TELOMERASE REGULATION IN *ARABIDOPSIS THALIANA*

A Dissertation

by

ANDREW DAVID LYLE NELSON

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2012

Major Subject: Biochemistry

Telomerase Regulation in *Arabidopsis thaliana*

TELOMERASE REGULATION IN *ARABIDOPSIS THALIANA*

A Dissertation

by

ANDREW DAVID LYLE NELSON

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Dorothy Shippen
Committee Members,	Gary Kunkel
	Jean-Philippe Pellois
	Alan Pepper
Head of Department,	Gregory Reinhart

August 2012

Major Subject: Biochemistry

ABSTRACT

Telomerase Regulation in *Arabidopsis thaliana*. (August 2012)

Andrew David Lyle Nelson, B.S., Southwestern Oklahoma State University

Chair of Advisory Committee: Dr. Dorothy E. Shippen

Telomeres form a nucleoprotein cap at the end of eukaryotic chromosomes. The telomere protein constituents repress the DNA damage response (DDR) and facilitate maintenance of terminal sequences by a specialized ribonucleoprotein complex called telomerase. In turn, factors involved in the DDR guarantee telomerase acts only in telomere homeostasis, and not at double-strand breaks (DSBs). Thus, the three pathways surrounding telomeres display incredible overlap and are immensely complex.

Here, I report a novel regulatory pathway that limits telomerase action during DNA damage. Duplication of the telomerase RNA subunit (TER) in Arabidopsis has given rise to a TER that is not required for telomere homeostasis. Indeed, this TER, termed TER2, is a competitive inhibitor of TER1 RNP complexes. Exposure to genotoxic agents results in TER2 upregulation and a subsequent inhibition of telomerase activity.

Using data from the 1,001 Arabidopsis genomes project, I determine that the TER duplication and inhibitory nature of TER2 is likely derived from a transposon-like element within TER2. This element is found throughout Brassicaceae, with at least 32 members in *Arabidopsis lyrata*. These findings highlight the complex and diverse mechanisms by which an organism will regulate telomerase action.

Here I characterize two members of the *A. thaliana* POT1 gene family. Contrary to POT1a, these proteins appear to have derived unique ways to perform their roles in chromosome-end protection. POT1b may protect telomeres as part of a TER2 telomerase RNP complex, as telomere defects only appear in the absence of both POT1b and TER2. POT1c is also appears to provide for chromosome end protection and appears to compete with POT1a to regulate telomerase access to the G-overhang. Together, these proteins represent part of a critical telomere capping complex distinct from CST.

Additionally, I describe a means for elucidating factors that regulate telomere addition at DSBs. This incredibly detrimental process, termed de novo telomere formation (DNTF), is toxic, and thus this work describes the first in depth characterization of DNTF in multicellular eukaryotes.

In summary, my work describes several novel regulatory and protective mechanisms for keeping telomeres and DSBs distinct.

DEDICATION

This dissertation is dedicated to my wife, Anna, who helps me be a better person, in all things, and in every way, and to my parents, for raising me right and supporting me all the way.

ACKNOWLEDGEMENTS

There are many special people responsible for helping me be where I am today. Science truly depends on the intellectual giants that precede us, and the ones that mentor us daily. First among these mentors is my PI, Dorothy Shippen. Dorothy has been a constant positive force, supporting my long and sometimes circutuous scientific journey to my doctorate. Dorothy's door is always open, and she is always willing to talk, dispencing advice on both scientific and personal topics. Dorothy has taught me to appreciate good and critical science, and to appreciate those who feel similarly. I will be eternally grateful to her for accepting me into her lab, where I feel that I have flourished.

I will also be eternally grateful to the faculty in the Biology and Chemistry departments at Southwestern Oklahoma State University, where I attained my Bachelors of Science. In particular, three professors immediately come to mind in preparing me for graduate school. These are Drs Peter Grant, Jason Johnson, and Muatassem Ubeidat. I was priviledged to work alongside the three of them, experiencing their (often humorous) take on science. All three professors made me feel like I was their colleage, and showed me that scientific pursuits are enjoyable.

I have had the joy of working with several extremely smart, nice, and fun scientists while in Dorothy's lab. Upon joining the lab, Yulia Surovtseva and Eugene Shakirov both mentored me and taught me everything they knew. I always enjoyed talking science with Eugene and Ross Warrington over a cup of coffee, even though Eugene is a Russian bastard (an inside joke that Eugene would appreciate)! The post-

docs Jon Lamb and Laurent Vespa were both integral members of the lab when I joined, and had an impact on my scientific career in different ways. Kalpana Kannan and Catherine Cifuentes-Rojas were both instrumental in characterizing the telomerase RNPs in Arabidopsis, and they helped me get my feet wet in characterizing regulators of telomerase. I will always be impressed by Xiangyu Song and Michelle Heacock's grasp of the scientific literature, as well as how nice and helpful they were. All of these people were members of the lab when I joined, and made me feel at home and welcome. I chose this lab, aside from the science, because of welcoming and collaborative attitude of everyone in the lab.

I have seen the addition of many new Shippen lab mates. I have enjoyed mentoring some of them, being mentored by others, and getting to know all of them. I have worked side-by-side with you all, and you are all great! You are all family to me, and I would do anything for you all. I wish every single one of you the best in the future, and look forward to more collaborations! I want to especially thank Katie for always being willing to help, even when she didn't really want to; Xiaoyuan and Callie for being sweet as well as smart (giving me hope for the next generation of the Shippen Lab); Kyle, Xintao and Hengyi for being such good sports when I felt like acting half my age (most of the time); Kara and Jung Ro for the scientific discussion. Kara and Kyle get a special nod for their political discussions (best not discuss libertarians when they are both in the room). Mark Beilstein will receive acknowledgements from me when I leave his lab, as this is my next stop on my scientific journey. Suffice it to say, while he was here, he made the lab fabulous!

I had the opportunity to mentor and work alongside a number of very talented undergraduates. The list is long, but a few stand out. Among these are Maren Cannel, Jessica Joseph, Amy Brinegar and Elena Douglass. I personally mentored Elena through her first year in the lab, and I feel privelidged to have had the chance. She is quick to learn, quick to laugh, and does things the right way, the first time. I know each and ever one of these special people will succeed at their life endeavors.

I would also like to take the time to thank the Biochemistry department, the faculty and staff, and my fellow graduate students. This department is the epitomy of family. From the moment I arrived, this department has welcomed me and given me every chance to learn and succeed. The BGA does an excellent job of integrating first year students into the fold, and Dr. Park and Pat Swigert help immensely in this fashion.

My committee, consisting of Drs Gary Kunkel, JP Pellois, and Alan Pepper put a great deal of thought and energy into my scientific progress. Like everyone else, they made me feel like a colleague. I would like to thank them all. JP deserves an extra mention, as he bore the brunt of my sense of humor (but it was mutual). JP, like Dorothy, was willing to give advice, both scientific and personal, whenever I needed it. Indeed, I felt like JP was an older brother (with less hair than by real brother).

I need to thank my family for all their support throughout my years in graduate school. My mom always thought I would be a veterinarian and care for animals, yet she never tried to dissuade me from graduate school. My parents instilled in me a desire to do the right thing, do a job the right way the first time, and take care of the ones I love. I would not be the person I am today if not for them.

To my friends, I can only say thanks, and that I am sorry. You have had to put up with a very silly person for a long time. I hope I have made it worth your time, and at the least, partially entertaining. Tim Devarenne and his lab have been especially welcoming to me (in part due to another special person). Joel, Julian, and Becki, aside from my lab mates, all hold a special place in my heart. Good luck, and best wishes. I hope to see you again down the road.

Finally, to the most important person in my life. Fortune and God's grace has blessed me with a loving and extremely intelligent partner in life. My wife Anna has been by my side since we graduate high school together. While life has not always been easy for us, Anna has always stuck by my side. She is extremely protective of me, and taught me to stand up for myself. We have grown up together, and have now earned our PhDs together. Also a Biochemistry PhD graduate student, Anna helped me navigate the travails of graduate school. As a sounding board and a beacon along the way, Anna has kept me focused and motivated. Thanks to you for all that you have done and continue to do.

I want to finish with a general reminder to all that I have known throughout graduate school. In some way, big or small, I am indebted to you for motivating me or helping me along the way. I fully believe that science thrives in an atmosphere such as the one in our department. In a way, this dissertation is part yours. Thank you all.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

In the late1930s, while America was suffering through the great depression and Europe was appeasing Nazi Germany, two geneticists were shedding light on the remarkable segments of DNA found at the end of eukaryotic chromosomes. Barbara McClintock, a maize cytogeneticist at the University of Missouri, followed the fate of broken chromosomes. McClintock observed a repair process at work on dicentric maize chromosomes after they broke apart during mitosis. McClintock termed this process "chromosome healing', since these chromosomes were safe from future fusion events (McClintock B, 1938). Around the same time, Herman Muller (University of Edinburgh), a geneticist studying the effects of x-rays on Drosophila chromosomes, found many genome rearrangements arising from the formation and repair of double-strand breaks (DSBs). Muller astutely noticed that chromosomal fusion did not arise from double strand breaks DSBs at the ends of the chromosomes (Muller, H.J. 1938). Working independently, both Muller and McClintock recognized that the ends of chromosomes had special properties that protected them from being covalently joined end-to end. Muller coined a name for these ends: telomeres (Muller, H.J. 1938).

Eventually scientists discovered that chromosomes are comprised of double stranded (ds) DNA, and that this DNA had a polarity to its structure. The inherent

---

This dissertation follows the style of *European Molecular Biology Organization Journal*.

makeup of the dsDNA helix complicates replication of the chromosome terminus, and led to theories on the End-Replication Problem, first laid out by Alexey Olovnikov in 1971 (Olovnikov A, 1971), and then by the renowned James Watson in 1972 (Watson J, 1972). Due to the semi-conservative mechanism of replicating linear, eukaryotic DNA, lagging strand synthesis results in a un-replicated segment of DNA each round of cell division. The loss of terminal DNA sequences would not be a problem in non-germline cells, as they would abide by Hayflick's limit (Hayflick and Moorhead, 1961), and stop dividing after ~50 cell divisions. However, some mechanism had to account for the complete transfer of full length chromosomes from one organism to its offspring.

It was not until the 1980s that Elizabeth Blackburn and Carol Greider solved this riddle, identifying an enzyme in the ciliated protozoan, *Tetrahymena thermophilla* that was capable of adding telomere repeats to a DNA molecule *in vitro* (Greider and Blackburn, 1985). Genetic and biochemical analysis followed, demonstrating that this enzyme, telomerase, consisted of a reverse transcriptase (TERT) and an associated RNA molecule (TER), both necessary for activity (Greider and Blackburn, 1987; Shippen-Lentz and Blackburn, 1990). Telomerase-mediated maintenance of telomere tracts has proven to be highly conserved across eukaryotes, and is crucial for cellular longevity. Indeed, altering the dynamics of telomere length maintenance, or perturbing the complex protein architecture that protects the telomeric DNA, has profound effects on integrity of the entire genome and on organismal viability (Jain and Cooper, 2010; Wellinger RJ, 2010; Boltz et al, 2012).

Telomeres and telomeric DNA

Elizabeth Blackburn was the first to sequence telomeric DNA. Her subject was mini-chromosomes encoding ribosomal RNA genes from Tetrahymena (Blackburn and Gall, 1978). Later, telomere sequence was acquired from a variety of related ciliates, revealing 20-70 copies of a G-rich repeat. Subsequently, linear plasmids abutted by the Tetrahymena telomere sequence were transformed into *Saccharomyces cerevisiae*, where they not only allowed for the maintenance of the plasmid, but were extended with the yeast telomere sequence (Szostack and Blackburn, 1982). This result suggested that there was some enzymatic mechanism capable of extending suitable G-rich substrates. It also suggested telomeres were a universal mechanism for safe-guarding the ends of chromosomes. This discovery spurred the search for the repeat sequence from other organisms, quickly revealing the G-rich nature of telomeres was conserved in multicellular eukaryotes (Figure 1-1A). Notably, across the metazoan lineage, the telomere repeat sequence is either TTAGGG (vertebrates), or TTAGG (insects), whereas the plantae lineage is dominated by the repeat TTTAGGG. The high conservation in telomere sequence suggests telomeres evolved as an early solution to genome maintenance in organisms with linear chromosomes.

While telomere sequence is highly conserved, telomere length varies across all eukaryotes, ranging from less than 50ntds in some ciliates to 150kb in tobacco (Murti and Prescott, 1999; Zellinger and Riha, 2007; Figure 1-1A). To maintain these varying lengths of telomeric DNA, mechanisms for maintenance and protection of chromosome ends have been established (see below).

A

Telomere length (kb)	Organism	Repeat Sequence
2-7	*Arabidopsis thaliana*	TTTAGGG
40-160	*Nicotiana tabacum*	TTTAGGG
2-40	Maize	TTTAGGG
~0.3	*S. pombe*	$TTACAG_{2-3}$
0.25-0.4	Tetrahymena	TTTTGGGG
2-30	Human	TTAGGG
20-150	*Mus musculus*	TTAGGG
-	Insecta*	TTAGG
0.3	*S. cerevisiae*	$G_{2-3}(TG)_{1-6}$

B

Fig 1-1. Telomere sequence conservation and general structure. (**A**) Telomere sequence and average length from several model systems. These numbers are adapted from Telomeres, Cold Spring Harbor Press, 2006. * This telomere sequence is not conserved among all insect species. (**B**) Schematic representation of the G-overhang in *Arabidopsis thaliana* (Riha et al, 2000).

G-overhangs and telomere secondary structure

The presence of a single-strand (ss) extension at telomeres was first recognized in the early 1980s, once again in ciliates (Klobutcher et al, 1981). Due to its G-rich nature, this extension was called the G-overhang. The G-overhang, now identified in all model organisms studied, including humans, ciliates, yeast, and Arabidopsis, ranges in length from 12 to 250 nt (Wellinger et al, 1993; Makarov et al, 1997; Wright et al, 1997; Henderson and Blackburn, 1989; Naduparambil et al, 2001; Riha et al, 2000) (Figure 1-1B). G-overhang length seems to be independent of overall telomere length, and varies throughout the cell cycle (Wellinger et al, 1993). A G-overhang would naturally form during lagging strand synthesis due to degradation of the RNA primer synthesized by polymerase α (Olovnikov, 1973). However, G-overhangs have been identified on both ends of chromosomes in humans (Lingner et al, 1995; Makarov et al, 1997). This finding argues that a resectioning event takes place on the leading strand, but how this occurs and the enzymes responsible are not well understood. Evidence from yeast demonstrates G-overhang formation during late S-phase, prior to, and independent of, telomerase action (Wellinger et al, 1993). In budding yeast, G-overhangs are present throughout the cell cycle, but a late S-phase increase in G-overhang length is necessary for proper telomere extension by telomerase (Larrivee et al, 2004). In mammals, G-overhangs range in length from 150-350nt, but unlike yeast, their size does not vary depending on stage of the cell cycle. Intriguingly, recent observations in Arabidopsis suggest that one end of the chromosome is blunt, in a Ku-dependent fashion (see below;

Karl Riha, personal communication). Thus, the architecture of the chromosome ends may not be completely parallel.

The G-overhang is an important aspect of telomeres, necessary for appropriate chromosome end protection. The G-overhang is also needed for the formation of appropriate telomere architecture, namely, the T-loop (Figure 1-2a). First observed in vertebrates in 1999, the T-loop has since been identified in cross-linked nuclear preparations from a host of organisms, including humans, mice, garden pea, and ciliates (Griffith et al, 1999; Cesare et al, 2003; Murti and Prescott, 1999). The T-loop is formed by the invasion of the G-overhang into duplex telomeric DNA, a process that is assisted by telomere binding proteins (TBP) *in vivo* (Griffith et al, 1999). A T-loop has also been observed in *Kluyveromyces lactis*, but not in the close relative *S. cerevisiae* (Cesare et al, 2008). Instead, in *S. cerevisiae*, the telomere is thought to form a simple fold-back structure. Formation of the fold-back is dependent on Rap1 (de Bruin et al, 2000) (Figure 1-2b). While the strategy chosen may differ among different organisms, these findings highlight the need to distinguish the end of the chromosome from a DSB.

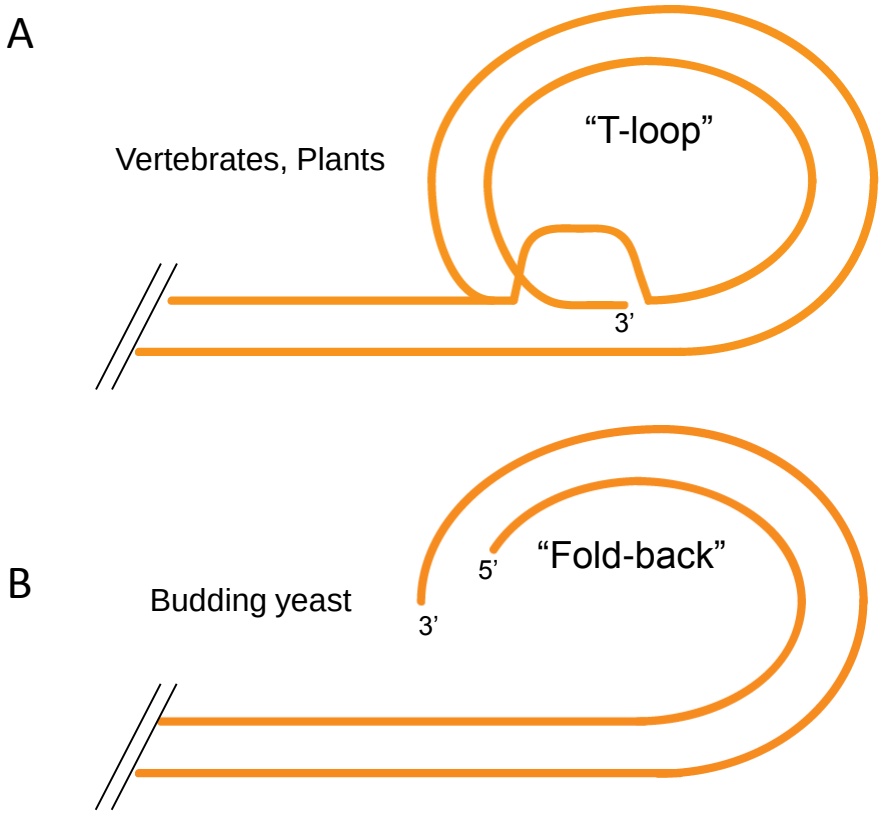


Fig 1-2. Secondary structure of telomeres. (**A**) T-loops have been visualized in both vertebrates and plants, and form through the G-strand invasion of upstream ds telomeric DNA (Griffith et al, 1999; Cesare et al, 2003). T-loops can form in vitro unaided, but in vivo, they form with the assistance of ds telomeric DNA binding proteins (primarily TRF2). (**B**) T-loops have not been identified in *S. cerevisiae*, but telomeres are believed to form “fold-back” structures to sequester the chromosome end (de Bruin et al, 2001).

The end replication problem and telomerase

When the end-replication problem was first described by Olovnikov and Watson, a model was proposed to explain the Hayflick limit (Hayflick and Moorhead, 1961; Olovnikov, 1971; Watson 1972). Gradual loss of telomeric DNA during each round of replication would result in cellular senescence and eventual death in cells lacking telomerase (Figure 1-3). However, Hayflick's limit was only observed in primary cell lines; single cell eukaryotes, such as Tetrahymena and yeast, and germline cells showed no such decline in population doubling over time (Figure 1-3).

This led Olovnikov to predict that these organisms/cell types had a means of countering telomere decline (Olovnikov, 1973). Olovnikov was correct. Twelve years later telomerase was purified from Tetrahymena (Greider and Blackburn, 1985), and then identified in other ciliates and in yeast (Shippen-Lentz and Blackburn, 1990; Counter et al, 1997). Although telomerase activities can be detected in human germ line and stem cells, it is absent in normal somatic tissue (Kim et al, 1994).

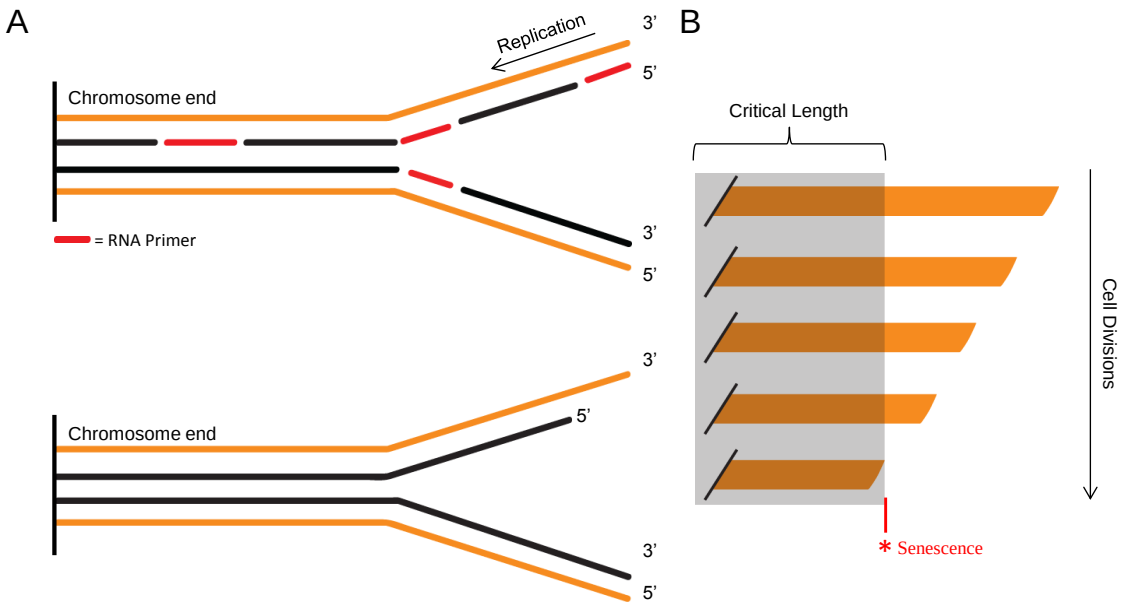


Fig 1-3. The end-replication problem and replicative senescence. (**A**) Lagging strand synthesis results in incomplete replication of the daughter strand (black), due to the inability to fill in the 5' end of the new daughter strand when the RNA primer is removed (Red). (**B**) Many cells follow a replicative senescence program. After a number of divisions, telomere attrition leads to a critical length, where senescence programs are activated (red *). Cells that manage to bypass this initial checkpoint either find a way to extend their telomeres by telomerase activation or recombination, or undergo massive chromosomal rearrangements.

Replicative senescence is now a well-established phenomenon in humans proposed to be important for tumor suppression (Kipling D, 1995). In humans, telomerase is repressed early in embryonic development in most tissues. Replicative senescence then arises from the progressive erosion of telomeric DNA due to incomplete replication, degradation, and recombination (Shay and Wright, 2010). Shortened telomeres are perceived as damaged DNA (see below) and activate DNA damage pathways (Schoeftner et al, 2009; Shore and Bianchi, 2009; Khadaroo et al, 2009). Shortened telomeres force the cell to exit the cell cycle and into the first stages of senescence (Toussaint et al, 2002). To avoid this fate, cells that must continue dividing submit their telomeres to telomerase for elongation. Thus, in stem and germ line cells, telomerase prevents replicative senescence by adding telomere repeats onto the G-rich strand.

The reverse transcriptase activity of telomerase lies in TERT (telomerase reverse transcriptase). TERT utilizes the telomere complementary sequence within its associated RNA subunit, TER, as a guide to bind to the G-overhang (Figure 1-4). Telomerase then reverse transcribes telomere repeats onto the telomere, using the G-overhang as a primer and TER as the template. Replication machinery, such as polymerases α and δ, coordinate with G-strand elongation by telomerase to fill in the C-strand late in S-phase (Ray et al, 2002; Price et al, 2010). This highly regulated process results in preferential elongation of the shortest telomeres to a species-specific setpoint, thus solving the end replication problem (Hug and Lingner, 2006).

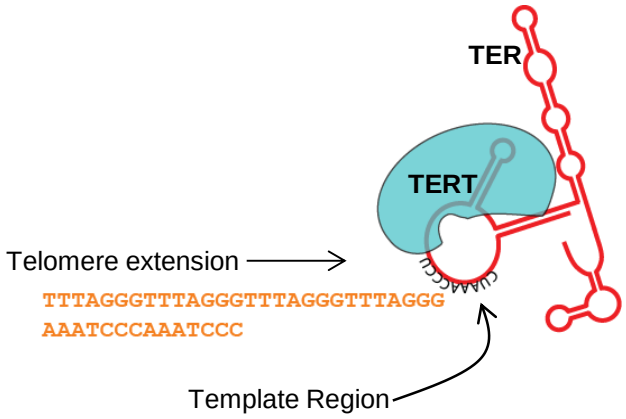


Fig 1-4. Telomerase is the primary mechanism for telomere elongation. Telomeres are maintained by telomerase, which is comprised of the reverse transcriptase TERT, and an RNA subunit TER. Telomerase aligns with the G-overhang, using the telomere complementary template region within TER (CUAAACCCU). Telomerase then reverse transcribes telomere repeats, using the telomere as a primer for nucleotide addition.

It should be noted that not all organisms regulate telomerase in the same fashion. Many single-celled eukaryotes restrict telomerase activity to a particular stage of the cell cycle (Gallardo et al, 2011; Bianchi and Shore, 2008). In mice, telomerase is active in most somatic tissue (Prowse and Greider, 1995), which may account for the increased incidence of tumors in these animals relative to humans. In Arabidopsis, as in humans, telomerase activity is present in actively dividing cells (young seedlings), but is gradually reduced as seedlings age and vegetative tissue fully develops (Forsyth et al, 2002; Fitzgerald et al, 1999; McKnight et al, 2002). Telomerase activity is most active in flowers and immature siliques, coinciding telomere elongation prior to meiosis and immediately afterward to generate extended homogeneous telomere tracts for all offspring (Fitzgerald et al, 1999).

TERT, the core enzymatic component of telomerase

TERT is a highly conserved reverse transcriptase that shares a common ancestor with the reverse transcriptase derived from Penelope-like retroelements (Gladyshev and Arkhipova, 2007). This observation supports the hypothesis that telomeres arose from retrotransposons. Indeed, retrotransposition at telomeres is still at work in Drosophila (Nosek et al, 2006). TERT ranges in size from ~66kD (*C. elegans*) to ~130kD in humans and Arabidopsis. Most TERT proteins contain a core RT domain, flanked by a long, sequence variable ~ 400 amino acid N-terminal extension, and a shorter ~150 amino acid C-terminal extension (Lingner et al, 1997; Harrington et al, 1997; Counter et al, 1997; Autexier and Lue, 2006). The minimal TERT from *C. elegans* retains only the RT

domain, this despite the observation that in most organisms, the N and C-terminal extensions (NTE and CTE) are required for activity (Malik et al, 2000). Telomerase is a processive enzyme in humans, Tetrahymena and likely Arabidopsis (Greider and Blackburn, 1985; Morin 1989; Fitzgerald et al 1999), implying steps of nucleotide addition followed by translocation and repositioning of the G-overhang with the RNA template. In contrast, the telomerase from yeast and mouse are non-processive. TERT makes multiple contacts with TER through an RNA binding domain within the NTE. The NTE binding sites on TER varies from species to species, perhaps due to the high sequence variability of the NTE (Autexier and Lue, 2006). This issue is also complicated by the paucity of TERs that have been identified.

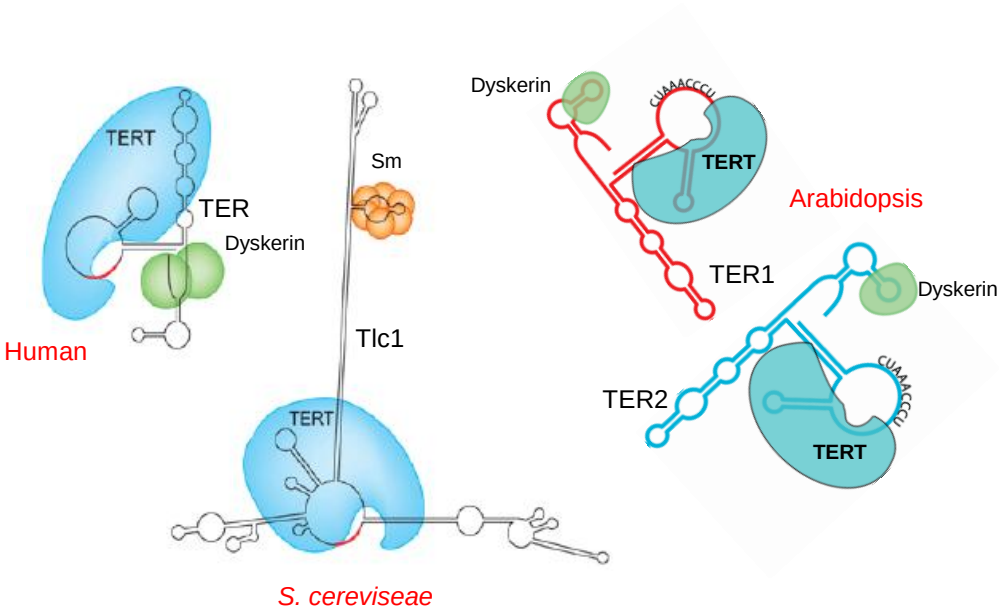


Fig 1-5 Minimal telomerase RNP complexes from select organisms. Human telomerase forms a dimer *in vivo* consisting of TERT, TER, and Dyskerin (Cohen et al, 2006). Yeast telomerase is also a dimer, consisting of the TERT, Tlc1 (TER), and the snRNP Sm proteins (Ly et al, 2003; Zappulla and Cech, 2004). Arabidopsis has at least two distinct telomerase complexes that may or may not dimerize (Cifuentes-Rojas et al, 2011). These complexes are formed from the two different TER molecules TER1 and TER2. Dyskerin associates with both TER1 and TER2 *in vivo* (Cifuentes-Rojas et al, 2011). Human and yeast complexes adapted from Teixeira and Gilson, 2007).

TER, the RNA component of telomerase

The essential nature of the RNA component in telomerase was shown when the addition of RNase to Tetrahymena extracts eliminated telomerase activity (Greider and Blackburn, 1987). Subsequent purification of Tetrahymena telomerase and characterization of co-purifying RNAs revealed TER (Greider and Blackburn, 1989). TER has been surprisingly difficult to identify in other eukaryotes because of its very extensive sequence divergence and a variety of different approaches have been used to find this moiety. TERs from ciliates, fission yeast, and Arabidopsis were identified through biochemical purification of telomerase and sequencing of co-purifying RNAs (Greider and Blackburn, 1989; Leonardi et al, 2008; Cifuentes-Rojas et al, 2011). (Figure 1-5). In contrast, budding yeast TER was identified by a screen designed to identify suppressors of telomeric silencing (Singer and Gottschling, 1994). Characterization of one of these suppressors revealed Tlc1. Mutations within the template region of this RNA confirmed its role in budding yeast telomerase (Singer and Gottschling, 1994). The human TER was identified by creating cDNA libraries from telomerase negative and positive cell lines, then subtracting the telomerase negative library from the positive library, enriching for genes involved in telomerase activity or regulation (Feng et al, 1995). Hybridization was then used to isolate cDNAs containing the predicted template sequence. Twelve RNAs were identified in this manner, one of which reconstituted telomerase activity.

One explanation for the rapid divergence of TER sequence is due to the constant need to evade cellular RNAi machinery (Neal Lue, personal communication). The

sequence divergence of TER within the Brassicaceae lineage will be discussed further below. In addition to sequence variability, TERs vary in length, from 150nt in Tetrahymena to ~1150nt in budding yeast (Greider and Blackburn, 1989; Singer and Gottschling, 1994). A mini-T yeast telomerase RNA of only ~150nt has been shown to reconstitute activity (Zappulla et al, 2005). Notably, the mini-T does not complement loss of full-length Tlc1 *in vivo* (Zappulla and Cech, 2006), supporting the conclusion that TER acts as a scaffold *in vivo* for the docking of multiple accessory proteins onto telomerase. Despite the drastic sequence dissimilarity and size differences, the secondary structure of TERs deduced through biochemical and phylogenetic analysis is highly conserved.

The most conserved feature of TER is the single-stranded template region, which usually corresponds to 1.5 copies of the telomere repeat. Not all organisms utilize the entire template, probably due to boundary elements on either side of the template region (Theimer and Feigon, 2006). These boundary elements constrain reverse transcription to just the template region, and also act as a spring by which the translocation step is initiated (Qiao and Cech, 2008). An additional conserved element in TER is a pseudoknot. The pseuodknot is a region of double and triple helical structures lying adjacent to the template (Theimer et al, 2005; Tzfati et al, 2003). The pseudoknot is believed to form a molecular switch, fluctuating between open and closed states to accommodate translocation along the G-overhang (Chen and Greider, 2005). The pseudoknot and template region is sufficient for robust activity *in vitro*, allowing for the high-resolution probing of this secondary structure (Qiao and Cech, 2008). More

importantly, the pseudoknot is essential for telomere maintenance by telomerase *in vivo* (Chen and Greider, 2004)

Within TER, there are several species-specific motifs lying outside of this core region. One region of interest is the H/ACA box, a motif present in small nucleolar RNAs that is necessary for the binding of the RNP maturation complex, Dyskerin (Chen and Greider, 2004; Fu and Collins, 2003). Dyskerin is essential for cell viability and functions to isomerize specific uridines into pseudouridines during the maturation process of ribosomal RNAs and spliceosomal snRNAs. Vertebrate TERs are not believed to undergo this conversion. Even so, the binding of Dyskerin is essential for human telomerase RNP biogenesis (Fu and Collins, 2003). The yeast homolog of Dyskerin, Cbf5p, has not been found to bind to Tlc1, but interestingly, the Arabidopsis Dyskerin, Nap57 (more below) has been found to bind the Arabidopsis TERs, suggesting an H/ACA box is present in both plant and animal TERs and furthermore that Dyskerin-mediated maturation of the telomerase RNP is conserved (Kannan et al, 2008).

Telomerase RNA processing

Expression and processing of TERs varies between species. In humans and yeast, TER (Tlc1) is transcribed by RNA Pol II, whereas in ciliates Pol III is responsible (Figure 1-6). The yeast telomerase RNA is processed similar to other Pol II transcripts in that it has a hypermethylated cap and is polyadenylated (Chapon et al, 1997; Bosoy et al, 2003). The 3’ end of Tlc1 may undergo two cleavage steps, one to remove the poly-A tail, and another that results in the removal of an additional 94 nts (Bosoy et al, 2003).

This final cleavage reaction is dependent on the yeast Nrd1-Nab3-Sen1 complex, which is also necessary for the maturation of other Pol II transcribed snoRNAs (Jamonnak et al, 2011; Ursic et al, 1997). In fission yeast, TERs with and without a poly-A tail have been identified indicative of a processing reaction (Figure 1-6) (Leonardi et al, 2008). After transcription, a complex of Sm proteins associates with the 3' end of TER, promoting a novel spliceosome-dependent cleavage reaction that corresponds to the first step of mRNA splicing (Box et al, 2008). This processing event causes release of Sm proteins and subsequent binding of the like-Sm (Lsm) complex to stabilize the mature 3' end. Lsm proteins remain associated with the telomerase RNP, protecting the 3' end from degradation (Tang et al, 2012).

The vertebrate TER contains a 5' hypermethylated cap, but does not acquire a poly-adenylated tail (Zaug et al, 1996). As with yeast TERs, ~100nts is removed from the 3' end of the RNA. In this case, the region required for 3' processing is located within the mature RNA (Fu and Collins, 2003). The H/ACA motif found at the processed 3' end, and the RNP proteins that bind it (Dyskerin et al) are most likely responsible for the correct processing of human TER (Zhang et al, 2011). As discussed in Chapter II, one of the two Arabidopsis TERs undergoes two novel processing events, one involving the cleavage of the TER2 3' end.

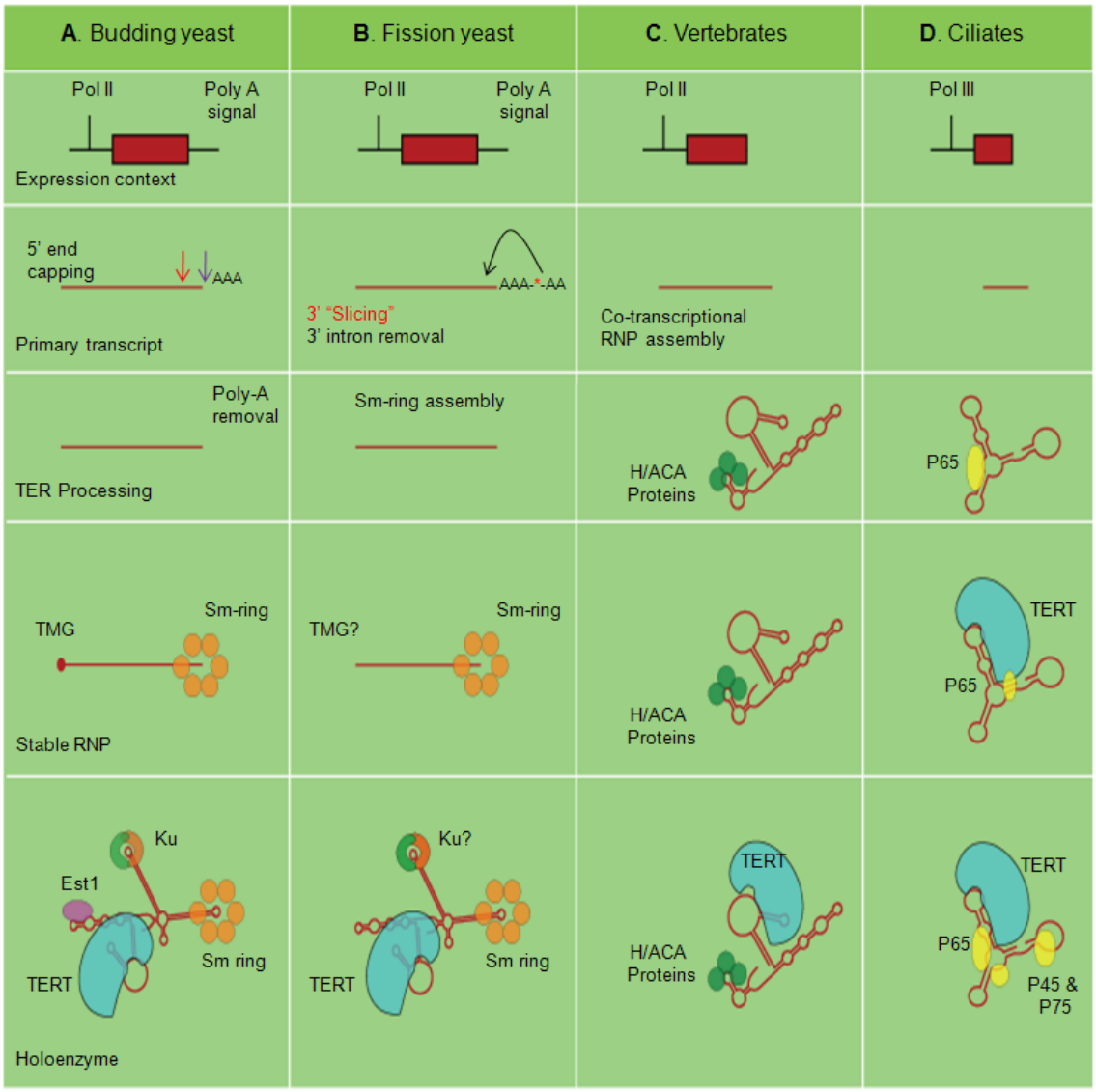


Fig 1-6. Biogenesis pathways for telomerase homoloenzymes. Schematic depicting processing steps of different TERs prior to holoenzyme formation. (**A**) Budding yeast Tlc1 is transcribed by Pol II and undergoes a polyadenylation step. Following polyadenylation, a 5' trimethylguanosine (TMG) cap is added, and snRNP Sm proteins associate. Two Tlc1 isoforms have been identified. One has undergone poly-A tract removal, while the other is missing an additional 94nt from the 3' end. (**B**) TER biogenesis in fission yeast occurs in a similar manner to budding yeast, in that it is Pol II transcribed and polyadenylated. However, the 3' end of the RNA is subjected to a novel "slicing" reaction, in which the downstream exon is not re-attached to the active, mature TER. This reaction is dependent on the spliceosome. (**C**) Vertebrate TER is transcribed by Pol II. However, RNP formation occurs in the nucleolus, involving binding of Dyskerin and H/ACA associated proteins. (**D**) Ciliate TERs are transcribed by Pol III. RNP biogenesis is primarily facilitated by p65, which stabilizes TER and promotes TERT association. This figure is modified from Collins K, 2006.

Telomerase-associated proteins

Telomerase activity can be reconstituted *in vitro* using only TERT and TER (Weinrich et al, 1997). However, *in vivo*, the telomerase holoenzyme contains several accessory factors that facilitate proper RNP biogenesis, enable optimal enzymatic activity, and negotiate recruitment to the chromosome end.

Telomerase accessory factors were first identified in *S. cerevisiae*, in a genetic screen for factors that stabilize linear chromosomes (Lundblad and Szostak, 1989; Lendvay et al, 1996). A series of mutants were identified that showed progressive telomere loss and were termed ever shorter telomere (EST). The first identified accessory factor was Est1p. Cells lacking Est1p do not immediately become inviable, but rather undergo gradual cellular senescence (Lundblad and Szostak, 1989). This genetic screen also identified Est2p, (TERT in budding yeast), Est3p, another telomerase accessory factor, and the telomere capping protein, Est4p/Cdc13p (Lendvay et al, 1996).

EST1 encodes an 82kD, highly basic protein that is capable of specifically binding telomeric ssDNA *in vitro* (Virta-Pearlman et al, 1996). Est1p also binds a bulged stem-loop in Tlc1, and this interaction is necessary for recruitment of the telomerase complex to telomeres (Seto et al, 2002). Telomerase recruitment to the telomere occurs through direct binding of Est1p to Cdc13p (see below; Lin and Zakian, 1995; Evans and Lundblad, 2002; Wu and Zakian, 2011). Est1p association with telomeres peaks during late S-phase, concomitant with the accumulation of other telomerase components (Fisher et al, 2004; Taggart et al, 2002; Tuzon et al, 2011). Est1p association with the chromosome end is dependent on its interactions with Cdc13p, as DNA binding by

Est1p has not been observed *in vivo*. Mutations that abolish the Est1p and Cdc13p interaction result in an EST phenotype (Evans and Lundblad, 2002). Est1p is reported to resolve G-quadruplexes *in vitro*, and notably, mutations that abolish this activity *in vitro* result in an EST phenotype *in vivo* (Zhang et al, 2010). These findings indicate that an Est1p-Cdc13p interaction is necessary for the initial recruitment of telomerase to the telomere, after which Est1p facilitates translocation of the enzyme along the G-overhang.

Putative *EST1* orthologs have been identified in various species of yeast, vertebrates, and plants. However, the essential telomere function of Est1p are confined to yeast. Est1 from fission yeast interacts with TER and the telomere capping protein Ccq1 (Beernink et al, 2003; Webb and Zakian, 2012). Like ScEST1, SpEST1 is necessary for telomerase recruitment (Moser et al, 2011; Beernink et al, 2003). Bioinformatic analysis revealed the putative EST1 ortholog in humans and plants (Reichenbach et al, 2003). Of the three EST1 genes in human s, only Est1a associates with telomerase. Over-expression of hEst1a induces uncapping of telomeres, resulting in telomere-telomere chromosomal fusions (Reichenbach et al, 2003). hEST1a is better known as SMG6, a critical component of the nonsense mediated RNA decay (NMD) pathway. All three human EST1 genes are implicated in this pathway (Chiu et al, 2003; Okada-Katsuhata et al, 2012). There are two EST1 genes in Arabidopsis (Riehs et al, 2008). Similar to humans, Arabidopsis homologs are involved in NMD and appear to be necessary for passage through meiosis. These genes do not have obvious telomere related functions (Riehs-Kearnan et al, 2012).

Another telomerase accessory factor was identified in the initial screen for budding yeast telomere mutants, Est3p (Lendvay et al, 1996). Est3p is a small protein, ~20kD, whose homo-dimerization is necessary for telomerase activity *in vivo* (Yang et al, 2006). Est3p does not require the presence of Tlc1 to interact with the TEN domain of Est2p (TERT). Indeed, this direct interaction is thought to stimulate telomerase beyond basal activity and potentially serve as an alternative means of recruitment to the telomere (Talley et al, 2011; Yen et al, 2011). The Est3p interaction with Est2p *in vivo* is conserved in other yeast species (Yen et al, 2011), but, no sequence or functional homologs have been identified outside of yeast.

While Est1p and Est3p regulate the telomerase complex in yeast, this function may have been transferred to the chromosome end in higher eukaryotes in the form of the telomere capping complex (see below). Relatively few proteins that assist telomerase in negotiating the chromosome end have been identified in multicellular organisms. Purification of endogenous telomerase complexes place their size between .65-2 Mda (Cohen et al, 2007; Schnapp et al, 1998; Cifuentes-Rojas et al, 2011). Even assuming a dimer of TERT, TER, and Dyskerin, this hypothetical complex barely reaches the lower observed MW. The size discrepancy in telomerase RNP complexes is a subject that will be covered in more depth later. However, known telomerase associated components in human cells are all implicated in RNP biogenesis, not in telomerase recruitment. Such factors include several Cajal body associated proteins, like Dyskerin, Dyskerin-associated proteins such as the ATPases Pontin and Reptin, and finally, TCAB1 (Venteicher et al, 2008; Venteicher et al, 2009). Cajal bodies are nucleolar compartments

where ribosomal, TER, and snRNP RNAs undergo biogenesis and maturation (Morris, 2008) Dykserin has been touched on briefly already, and will be addressed further in Appendix (App-I). The ATPases Pontin and Reptin physically interact with one another, have roles in chromatin remodeling, and interact with small nucleolar RNA complexes, in a fashion similar to Dyskerin (Newman et al, 2000; Rottbauer et al, 2002). Pontin and Reptin most likely associate with telomerase through Dyskerin. Knockdown of either protein results in reduced telomerase activity *in vivo* due to loss of TER accumulation (Venteicher et al, 2008). While Pontin and Reptin are most likely necessary for accumulation and biogenesis of the mature telomerase molecule, they are not necessary once the telomerase RNP has assembled (Venteicher et al, 2009).

TCAB was identified in the same type of purification scheme that recovered Pontin and Reptin (Venteicher et al, 2008). TCAB also physically associates with Dyskerin, but requires TER to associate with TERT, suggesting it is in a complex situated at the H/ACA box along with Dyskerin. TCAB1, unlike Pontin and Reptin, associates with an active telomerase holoenzyme (Venteicher et al, 2009; Zhong et al, 2011). Interestingly, TCAB1 localizes to Cajal bodies, where Dyskerin and TER associate in human cells (Venteicher and Artandi, 2009). TCAB1 depletion does not affect overall activity levels, but it does affect telomerase localization to Cajal bodies. Cajal bodies are believed to store telomerase until the S-phase of the cell cycle, whereupon telomerase is released to engage the telomere (Venteicher and Artandi, 2009). TCAB1 depletion decreases telomerase association with the chromosome end during S-phase and results in progressive telomere shortening (Venteicher et al, 2009;

Zhong et al, 2011; Batista et al, 2011). The mechanism of TCAB1 function is still unknown, but it provides an interesting window into regulation of the telomerase holoenzyme in humans. The emerging view of telomerase accessory proteins in Arabidopsis will be discussed below.

Telomere-associated proteins

Telomeres must be sequestered into complex amalgams of protein and DNA to prevent their identification as DSBs. Thus, telomere capping involves physically obscuring telomeric DNA, and the active repression of DDR pathways. In addition, the telomere capping proteins must regulate access of telomerase to the G-overhang (See below). While a few of the telomere capping factors are species-specific, there is more conservation at the chromosome end than appear to be in the telomerase complex. The protein composition at telomeres can be broadly divided into ds and ss telomeric DNA binding proteins and the factors that bridge these two domains.

In yeast, ds telomeric DNA is associated with the essential transcriptional repressor/activator protein, Rap1 (Longtine et al, 1989; Runge and Zakian, 1990; Lustig et al, 1990). Rap1 contains a Myb DNA binding domain, which shows flexible specificity for a repeat sequence found within the heterogeneous telomere tracts (Konig et al, 1996; Longtine et al, 1989). Myb DNA binding domains are common in the ds telomeric DNA binding proteins, and will be discussed in the context of the TRF proteins below.

Over-expression of Rap1 leads to increased telomere length and heterogeneity, partly a result of increased telomere recombination (Conrad et al, 1990). Rap1 is a negative regulator of telomerase thought to function in a "counting" mechanism to control telomere length (Marcand et al, 1997; Shore and Bianchi, 2009). Long telomeres are bound by more Rap1, leading to decreased telomerase action (Levy and Blackburn, 2004). Telomerase inhibition may be due to simple telomere inaccessibility (Teixeira et al, 2004). High Rap1 concentration at telomeres is necessary to fold telomeres back into a telomerase non-extendible state (de Bruin et al, 2000; Teixeira et al, 2004). This Rap1 telomere counting mechanism signals through the conserved DNA damage response (DDR) kinases Tel1/Mec1 (ATM/ATR) (Ray et al, 1999; Craven et al, 1999). Rap1 also promotes telomere silencing through interactions with the chromatin remodelers and silencers Sir3 and Sir4 (Hardy et al, 1992; Bourns et al, 1998). Rap1 mediates telomere length homeostasis and suppression of the DDR through interactions with two Rap1 interacting factors, Rif1 and Rif2 (Wotton and Shore, 1997; Levy and Blackburn, 2004; Teixeira et al, 2004).

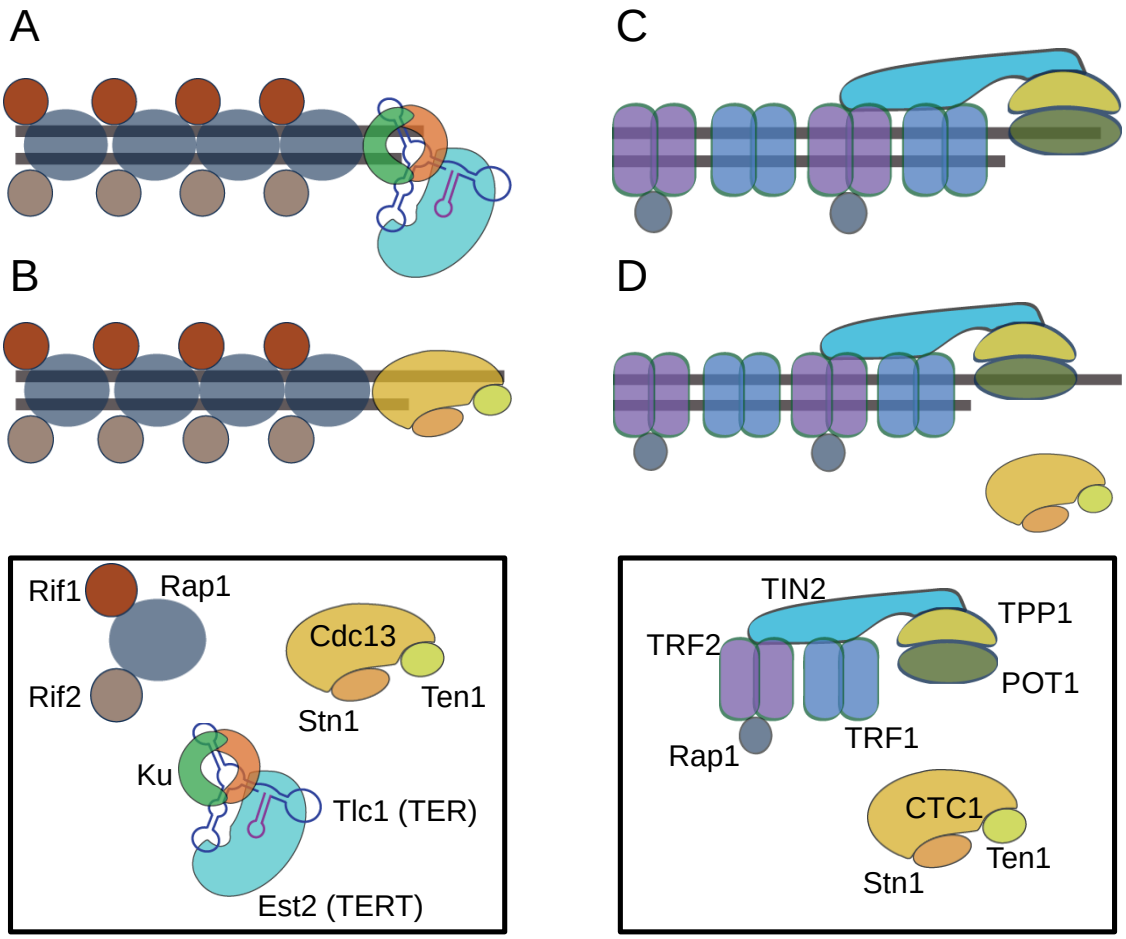


Fig 1-7. Telomere capping complexes in budding yeast and vertebrates. (**A**) Potential telomere capping complex during the G_1 phase of the cell cycle in budding yeast. The predominant dsDNA binding telomere protein is Rap1, which also associates with two other proteins, Rif1 and Rif2. G-overhangs are very short in budding yeast outside of S-phase, so there is the potential for no CST (Cdc13/Stn1/Ten1) binding. Instead, telomere ends could be bound by a complex of Ku and an inactive telomerase (proposed model). (**B**) During S-phase, the G-overhang's elongate, and Cdc13, Est1, and TERT are enriched, coinciding with telomere elongation and replication. (**C**) In vertebrates, telomeres are primarily bound by the six-membered complex called shelterin. The dsDNA binding proteins consist of TRF1 and TRF2, which homodimerizes. TIN2 serves as a bridge between the TRF proteins and the ss strand binding heterodimer, TPP1/POT1. In vertebrates, there is always at least a short G-overhang (~40nt) that gets extended preceding telomerase activity. (**D**) With the identification of a CST complex in vertebrates, the possibility for overlapping functions with yeast CST arises. One possible common role for CST is to help initiate replication of the telomere during S-phase.

The ds telomere region in vertebrates is bound by components of a conserved six-membered complex termed shelterin. Shelterin is composed of two dsDNA binding proteins, TRF1 and TRF2, a ssDNA binding protein POT1, TRF1 and TRF2 interacting nuclear protein TIN2, Rap1, and the POT1 interacting protein TPP1 (Figure 1-7). All six components can bind one another in the absence of DNA, but show remarkable specificity for the chromosome terminus (Palm and de Lange, 2008).

A factor showing high specificity to the human telomere repeat was first identified using the human telomere repeat as "bait" in an EMSA experiment involving nuclear extracts (Zhong et al, 1992). Further purification and mass spec analysis identified this was a single protein, TRF1, which specifically co-localized with telomeres (Chong et al, 1995). TRF2 was first identified based on homology to TRF1 and initial characterization showed TRF2 bound ds telomere repeats (Bilaud et al, 1996; Bilaud et al, 1997; Broccoli et al, 1997). TRF1 and TRF2 are highly abundant proteins that display homodimerization through a highly conserved dimerization domain (Bianchi et al, 1997). TRF1 and TRF2 also contain a myb DNA binding domain, similar to the ScRap1 DNA binding domain. Whereas TRF1 is ~60kD and slightly acidic at neutral pH, TRF2 is 67kD and highly basic at cellular pH (Chong et al, 1995; Broccoli et al, 1997). TRF2 shows reduced DNA binding *in vitro*, but this does not impair TRF localization to telomeres *in vivo*. Both TRF1 and TRF2, like Rap1 in budding yeast, are negative regulators of telomere length (Smogorzewska et al, 2000). TRF2 is believed to perform this role by promoting formation of t-loops, a telomeric substrate unsuitable for telomerase (Smogorzewska et al, 2000). TRF2 is capable of creating positive supercoils

that unwind DNA and promote strand-invasion (Amiard et al, 2007). In addition, TRF2 contains a Gly/Arg-rich domain that shows specificity for Holliday junctions, which would be present at a t-loop (Amiard et al, 2007; Fouche et al, 2006). Both TRF1 and TRF2 change the secondary structure of telomeric DNA, creating loops of DNA around themselves, acting like telomere specific histones (Palm and de Lange, 2008).

Rap1 in vertebrates is a TRF2 binding partner and is dependent on TRF2 for telomere localization, as it does not show DNA binding capabilities of its own (Celli and de Lange, 2005). The function of Rap1 is poorly understood, but recent data suggests that it is necessary for the inhibition of NHEJ at telomeres (Sarthy et al, 2009). TIN2 is a critical component of shelterin, that bridges between the ds and ssDNA binding proteins, TRF1/TRF2 and TPP1 (Kim et al, 2004; Houghtaling et al, 2004). Loss of TIN2 has a profound effect on shelterin stability (Ye et al, 2004). In summary, the ds telomere binding proteins in both yeast and vertebrates are necessary to regulate telomere length, telomerase access to the telomere, and prevent an illegitimate DDR (Denchi EL, 2009).

The occlusion of the G-overhang is the responsibility of two different conserved complexes. The first complex, CST, is a heterotrimer comprised of Cdc13, Stn1, and Ten1. First identified in budding yeast, this complex was originally thought restricted to the family Saccharomycetaceae, which includes the model organisms *S. cerevisiae*, and *C. albicans*. However, it is now clear that this complex is functionally conserved in multicellular eukaryotes (Martin et al, 2007; Song et al, 2008; Surovtseva et al 2009, Miyake et al, 2009).

Budding yeast Cdc13 is a multifunctional 100kD protein that harbors multiple OB (Oligosaccharide/Oligonucleotide Binding Motif) folds within its N-terminal DNA binding domain. This OB fold is necessary for binding to an 11-nucleotide sequence found within the G-overhang that is complementary to the template region within Tlc1 (Mitton-Fry et al, 2002). *In vitro* DNA binding experiments revealed that Cdc13 has high affinity for ss telomeric DNA substrates, and naturally obscures the G-overhang from template binding by telomerase (Zappulla et al, 2009). Cdc13 binds to the telomerase accessory factor Est1p (Qi and Zakian, 2000). Mutations within the Est1 binding domain on Cdc13 also show an EST phenotype (Evans and Lundblad, 2002). Telomerase recruitment via the Est1-Cdc13 interaction is regulated in an ATM/ATR dependent fashion (Tseng et al, 2006), which will be discussed below.

An essential role for Cdc13 is to protect the telomere from exonucleolytic degradation and to prevent the terminus from being recognized as a DSB (Nugent et al, 1996). Temperature-sensitive Cdc13 mutants that impair these functions result in C-strand resection, activation of a massive DNA damage response, and subsequent cell cycle arrest (Garvik et al, 1995; Nugent et al, 1996, Lin et al, 1996). Cdc13 functions in concert with its binding partners, Stn1 and Ten1. Stn1 was originally identified as a suppressor of a temperature sensitive Cdc13 mutant (Grandin et al, 1997). Stn1 and Cdc13 physically interact both *in vitro* and *in vivo*, and yeast mutant for Cdc13, but over-expressing Stn1, are viable. Ten1 was identified in a screen to identify suppressors of the Stn1-13 mutant allele, which results in abnormally long telomeres (Grandin et al, 1997, Grandin et al, 2001). Ten1 interacts with both Cdc13 and Stn1. While Stn1 is

believed to assist Cdc13 in controlling telomerase action, Ten1 is necessary for Cdc13's suppression of exonucleolytic degradation (Xu et al, 2009). Loss of any of the CST components results in the same phenotype, indicating these three components function together to cap the chromosome end.

Little is known about the role of CST (CTC1/STN1/TEN1) role at vertebrate telomeres (Surovtseva et al 2009, Miyake et al, 2009). hCST binds to telomeric ssDNA as a complex, but individual components display little affinity for DNA (Miyake et al, 2009). Knockdown experiments of these components demonstrate that the CST complex protects telomeres independent of the other ssDNA binding complex in vertebrates (POT1/TPP1, see below). An important clue for the function of CST comes from the discovery that CTC1 and STN1 are components of the DNA polymerase-α complex, and hence are believed to facilitate replication of the chromosome terminus (Casteel et al, 2009). Mutations within CTC1 result in shortened telomeres and DNA damage loci (Anderson et al, 2012) and have recently been identified as the cause of several genetic diseases (Anderson et al, 2012; Polvi et al, 2012). STN1 binds to shelterin components, and removal of this interaction domain leads to abnormal telomere elongation (Wan et al, 2009). Thus, in vertebrates, both CST and shelterin are essential for complete telomere protection and replication.

The other ss telomere DNA specific complex conserved in many eukaryotes was first identified in the ciliate *Oxytricha nova* (Gottschling and Zakian, 1986; Price and Cech, 1987; Gray et al, 1991). This heterodimeric complex, consisting of TEBP α/β, shows high affinity for ss telomeric DNA. The α subunit directly binds DNA, whereas

the β subunit confers tighter binding for the complex *in vivo,* but does not bind DNA itself (Gray et al, 1991). Orthologs of TEBPα/β have been identified in vertebrates and fission yeast (Bauman and Cech, 2001; Liu et al, 2004; Miyoshi et al, 2008). A bioinformatics search for TEBP-α like proteins revealed a potential homolog, referred to as Protection Of Telomeres 1 (POT1) (Baumann and Cech, 2001). POT1, like TEBP-α, shows high specificity for ss telomeric DNA. TPP1, the functional human homolog of TEBP-β, was identified due to its strong interaction with POT1 (Liu et al, 2004). Similar to Oxytricha TEBP-α/β, TPP1 and POT1 dimerization allows for tighter binding to ssDNA (Lei et al, 2004; Xin et al, 2007; Wang et al, 2007). Unlike TEBPα/β, human binding activity is not limited to the extreme 3' end of the G-overhang. In accordance with the longer G-overhangs in humans (50-300nt), POT1/TPP1 dimers can coat the G-overhang, sequestering it away from harmful activities (Taylor et al, 2011).

TPP1 binds to TIN2, linking POT1/TPP1 to the ds portion of shelterin (Takai et al, 2011). TIN2 tethering is proposed to increase the odds of POT1 out-competing RPA for ssDNA binding, thus inhibiting ATR activation (see below). Indeed, POT1 mutants unable to bind DNA still localize to telomeric DNA and co-purify with shelterin components (Liu et al, 2004). The interaction between shelterin and TPP1/POT1 is critical for telomere length maintenance and inhibition of the DDR (Ye et al, 2004; Liu et al, 2004; Hockemeyer et al, 2007).

TPP1 also displays weak interactions with TERT, and gel filtration profiles show TPP1 co-fractionation with TERT in a ~2Mda complex, suggesting earlier estimates of telomerase complex size may have included TPP1/POT1 (Xin et al, 2007). The

TPP1/POT1 complex increases telomerase processivity on ssDNA *in vitro* (Wang et al, 2007). This observation indicates that POT1/TPP1 increases telomerase activity on G-overhangs long enough to accommodate both telomerase and POT1/TPP1.

Just as the predominant dsDNA binding components share the myb-DNA binding domain, all known ss telomeric DNA binding proteins display tight and specific binding through an OB fold. POT1/TPP1, Cdc13/Stn1/Ten1 and the corresponding functional homologs from other organisms contain one or multiple OB folds that contact DNA (Baumann and Cech, 2001; Xin et al, 2007; Theobald and Wuttke, 2004). In the case of human and fission yeast POT1, crystal structure data suggests that DNA binding is primarily mediated through the first OB fold, whereas a second OB fold is necessary for stabilization of the substrate (Baumann and Cech, 2001).

In conjunction with binding DNA, OB folds also participate in RNA binding and protein-protein interactions (Murzin AG, 1993; Theobald et al, 2003). OB folds are ancient structural motifs typified by a five-stranded β-barrel in a 1-2-3-5-4-1 arrangement, with α-helices separating strands 3 and 4, and loops between the strands making primary contacts with DNA (Figure 1-8) (Kelly et al, 1998; Raghunathan et al, 1997; Bycroft et al, 1997). Most OB folds can be readily predicted, such as the two 140aa OB folds present in the N-terminus of all known POT1 proteins.

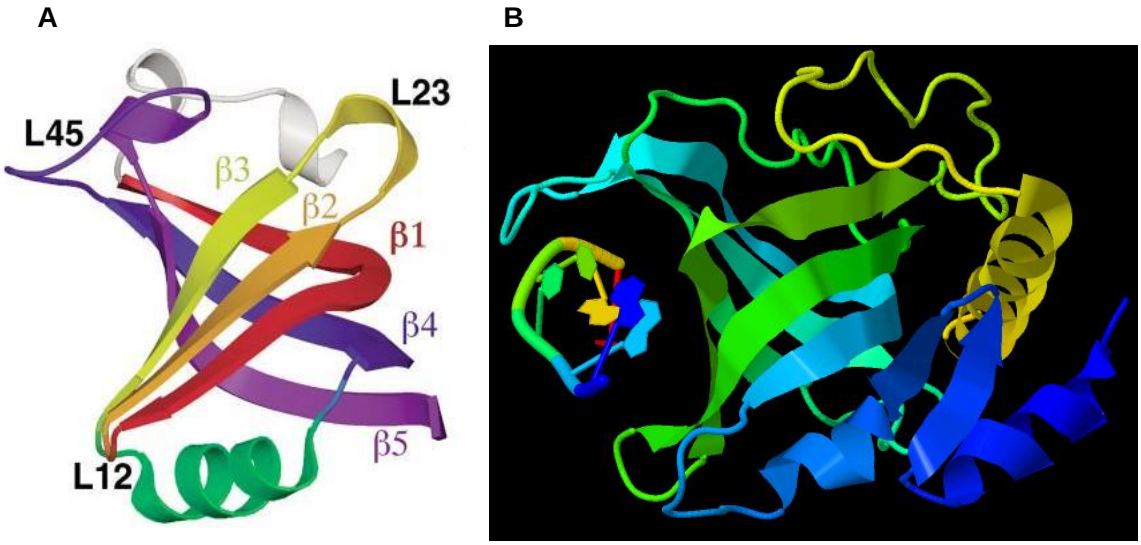


Fig 1-8. The Oligonucleotide-Oligosaccharide binding (OB) fold. (**A**) Ribbon diagram of an OB-fold. OB-fold domains are characterized by two three-stranded antiparallel β-sheets, where β1 takes part in both sheets. (**B**) Ribbon diagram of the first OB-fold of SpPOT1 bound to ssDNA. ssDNA fits into the groove created by L45 and L12. Extensive hydrogen bonding and base-stacking occurs in this groove to stabilize the ssDNA substrate.

Substrate specificity is a rarity for OB-folds (Bochkarev and Bochkarev, 2004; Krapp et al, 1998), which makes telomere specific OB-fold proteins all the more puzzling. The RPA trimer is a perfect example of OB-fold containing proteins that contact DNA non-specifically. RPA consists of six OB-folds, four of which bind DNA (Bochkarev and Bochkarev, 2004). RPA binds ssDNA irrespective of sequence, and will bind telomeric DNA in the absence of POT1 and initiate a DDR (Flynn et al, 2011). Thus, a highly conserved function for the POT1/TPP1 dimer is to prevent RPA binding and downstream ATR activation (Churikov and Price, 2008; Palm et al, 2009; Hockemeyer et al, 2006; Flynn et al, 2012).

A less understood component of telomeres is the non-homologous end-joining component Ku. Ku is a ubiquitous and highly conserved heterodimer comprised of Ku70 and Ku80. Ku is more famously known for its role in DNA damage repair. It binds to the ends of duplex DNA irrespective of sequence, helping to stabilize these ends during the repair process (Moore et al, 1996). With this in mind, the need for Ku to stabilize telomeres is counterintuitive (Boulton and Jackson, 1996; Nugent et al, 1998). The Ku heterodimer can perform different roles at DSBs and telomeres (Ribes-Zamora, 2007). ScKu is necessary for the formation of heterochromatin at telomeres, which leads to transcriptional silencing (Nugent et al, 1998, Bertuch et al, 2003). ScKu is also responsible for prevention of telomere recombination, C-strand resection (See below)

Interestingly, ScKu is important for the recruitment of telomerase to yeast in a Tlc1-dependent, Est1-independent fashion (Nugent et al, 1998; Stellwagen etl al, 2003; Fisher et al, 2004). This recruitment is responsible for anchoring telomeres at the nuclear

periphery in yeast (Schober et al, 2009). In both budding yeast and vertebrates, Ku interacts with telomerase via a 48nt hairpin in TER (Stellwagen et al, 2003; Ting et al, 2005). This interaction is responsible for recruiting an inactive form of yeast telomerase to the G-overhang during G1 of the cell cycle (Stellwagen et al, 2003; Fisher et al, 2005). One possible explanation for Ku-mediated telomerase recruitment comes from the yeast *Candida albicans*. In this system, telomerase is one of many components necessary for telomere capping. Absence of either TERT or TER results in increased C-strand resection, a role that is most likely dependent on Ku recruitment (Hsu et al, 2007). Additionally, the Ku/Tlc1 interaction is also necessary for the recruitment of telomerase to the site of DSBs (Stellwagen et al, 2003), where telomere repeats are then added. Why a protein that is necessary for healing of DSBs would be involved in such a detrimental process is quite interesting, and will be discussed below.

Telomeres and DNA damage repair machinery: the enemy at the gate or the one within

The end of a eukaryotic chromosome must be hidden from two competing forces. One of these is exonucleolytic activity, and the other is unwanted DNA repair that culminates in one chromosome end covalently attaching to another. These forces are prevented by the constant vigilance of telomere capping proteins. However, telomere replication and extension provide a point of vulnerability. Telomere lengthening and replication are tightly linked. Indeed, telomere extension follows on the heels of replication forks progression through or from the telomere end (Marcand et al, 2000;

Dione and Wellinger, 1998; Bianchi and Shore, 2007). Following DNA replication, a new complex of proteins must assemble on the telomeres of each new chromosome. The issue gets even more complicated, because much of the DNA damage repair machinery is already present at telomeres, performing essential roles in telomere maintenance (see below). If telomeres are ever unprotected, DDR proteins switch roles immediately, initiating repair processes that can result in degradation of terminal sequences.

Cells are alerted to DNA damage through the activities of two key protein kinases, ATM and ATR. ATM primarily responds to dsDNA breaks by the break-binding complex Mre11 (Khanna and Jackson, 2001; Riha et al, 2006), whereas ATR senses accumulation of ssDNA through the activities of the ssDNA binding protein heterotrimer RPA (Shiloh Y, 2003). Upon sensing damaged DNA, these kinases phosphorylate the histone variant H2AX in a large region surrounding the site of damage (Kinner et al, 2008; Amiard et al, 2010; Szilard et al, 2010). Additionally, ATM and ATR activate kinases responsible for blocking cell cycle progression until the DNA has been repaired. In order to repair DNA breaks, ATM and ATR recruit appropriate DNA repair machinery.

Of the two main DNA repair pathways, homologous recombination (HR) and non-homologous end-joining (NHEJ), the latter is strongly preferred in multicellular eukaryotes (Riha et al, 2006; Symington and Gautier, 2011). HR is limited to S and G2, where sister chromosomes are present to act as a template for repair. In contrast, NHEJ occurs during G1. There are two NHEJ pathways, classical NHEJ (C-NHEJ), and alternative NHEJ (A-NHEJ) (Stracker and Petrini, 2011). C-NHEJ is a robust DNA

repair pathway that begins with Ku-mediated stabilization of the DNA around the break site (Riha et al, 2006). Ku is a ubiquitous protein that associates with both telomeres as well as dsDNA breaks, and prevents excessive exonucleolytic degradation (Symington and Gautier, 2011). At dsDNA breaks Ku recruits factors necessary for the processing and re-ligation of the DNA (Boulton and Jackson, 1996). These factors include the exonuclease Artemis, which processes ends to create a suitable substrate for DNA ligase IV (Moshous et al, 2001). Ku is also capable of recruiting polymerases necessary for filling in gaps that arise during re-alignment of the resected DNA (Ma et al, 2004). A-NHEJ, or microhomology mediated NHEJ, centers around the complex Mre11, Rad50, and NBS1/Xrs2 (MRN/X) (McVey and Lee, 2008; Symington and Gautier, 2011). MRN/X, like Ku, stabilizes dsDNA breaks, then initiates the resectioning of DNA while scanning for small regions of homology between the ends being repaired (Stracker and Petrini, 2011).

Components of both NHEJ pathways are present at telomeres, including the kinases responsible for activating them, ATM and ATR (Palm and de Lange, 2008). Ku, as mentioned above, is responsible for stabilizing telomere ends in humans, yeast, and Arabidopsis (Wang et al, 2008; Bertuch and Lundblad, 2003; Riha and Shippen, 2003). ATM and ATR both seem to regulate telomerase activity, and repress the DDR at telomeres (Moser et al, 2011; Tseng et al, 2009). MRN/X is also associated with telomeres, where it is important for C-strand resection (Williams et al, 2010). In budding yeast, ATM association with telomeres is MRN/X dependent, while MRN/X activity is ATM dependent (Martina et al, 2012). Interestingly, these enzymatic activities are

necessary for proper telomere lengthening, most likely by creating a suitable binding site for telomerase (Lamarche et al, 2010; Bonetti et al, 2010). In fission yeast, ATM and ATR kinase activity is essential for recruitment and activation of telomerase (Yamazaki et al, 2012). Ku, through its interactions with Tlc1, is necessary for the proper accumulation of telomerase at telomeres in G1, perhaps to form a telomere protective complex (Stellwagen et al, 2003; Peterson et al, 2001). Thus, the function of Ku at telomeres is paradoxical. The structure of the chromosome end largely resembles substrates for the DDR machinery, and yet this function must be subverted specifically at telomeres and nowhere else in the genome. How Ku distinguishes between DSBs and natural telomeres is discussed below.

De novo telomere formation

One interesting facet of telomerase enzymology is its promiscuous nature. *In vitro* telomerase activity assays can be performed with a primer terminating in a single G residue. How then is telomerase prevented from acting at a DSB? The addition of telomeric sequence at an interstitial break site results in the loss of the centromere-distal chromosome DNA fragment. This outcome results in partial monosomy, and is the source of several genetic diseases such as Phelan/McDermid syndrome and alpha thalassemia (Bonaglia et al, 2011; Wilkie et al, 1990). McClintock observed de novo telomere formation, often referred to as “chromosome healing”, while characterizing the recovery of fused chromosomes in maize (McClintock B, 1941). Dicentric chromosomes break during anaphase, and this breakage-fusion-bridge cycle continues until a stable

chromosome is formed, or the cell enters senescence (Riha et al, 2006). McClintock observed that broken chromosomes could be repaired in the germline, by some mechanism that prevented the chromosomes from re-fusing (McClintock B, 1941). Of course this mechanism is telomerase, and thus, the recruitment of telomerase to chromosome breaks is not quite so peculiar. Given the choice between the loss of a cell or the loss of part of a chromosome, a mechanism has been developed to retain the cell. In chapters II and VI of this dissertation, experimental data are presented that build on the body of work from budding yeast to further our understanding of the mechanism for telomerase recruitment and regulation at dsDNA breaks.

Telomerase regulation during DNA damage

DSBs are an incredibly toxic form of DNA lesions arising from mistakes in DNA replication, genotoxic agents such as reactive oxygen species, and ionizing radiation (Xu and Price, 2011). As touched upon above, HR and NHEJ pathways are employed when such DSBs arise. However, mistakes happen, and a potential outcome of failed DSB repair is *de novo* telomere formation (DNTF). Analysis of sites of DNTF in human and murine systems reveal that very little sequence is lost at the break site, in line with Ku's role at stabilizing and preventing resection at DSBs (Sprung et al, 1999; Hannes et al, 2010). In budding yeast, telomerase can add telomere repeats at DSBs with little or no telomere repeat homology (Figure 1-8B) (Stellwagen et al, 2003). DNTF is dependent on an interaction between Ku and Tlc1, placing telomere addition in direct competition with DSB repair machinery. Intriguingly, Ku binds TER in humans, and in chapter II we

present data demonstrating Ku binding to Arabidopsis TER2 (Ting et al, 2005). In addition, data from Arabidopsis (Chapter VI) suggests that the requirement for Ku in DNTF is functionally conserved, raising the question of why the recruitment of telomerase to sites of DNA damage is conserved.

One possibility is that a Ku-mediated DNTF pathway reflects Ku's ability to recruit telomerase to natural telomeres. In this scenario, recruitment of telomerase to sites of DSBs is unintentional, arguing that mechanisms would be in place to prevent this from occurring. First I will discuss recent data indicating the constituents of the Ku heterodimer, Ku70 and Ku80, enable the complex to perform different functions, depending on the orientation of loading onto a DNA substrate.

Targeted mutagenesis on Ku70 and Ku80 in budding yeast revealed that telomere silencing is mediated by a specific domain on Ku80, while NHEJ requires a specific domain on Ku70. Thus, a "two-face" model has been proposed (Figure 1-8) (Ribes-Zamora et al, 2007). When Ku loads onto dsDNA, Ku80 faces inwards, away from the DNA end, and Ku70 faces outwards (Lopez et al, 2011). Ku80's orientation at telomeres activates telomere spreading through interactions with the Sir family of proteins (Boulton and Jackson, 1998). Ku80 is also the site of interaction with TER (Stellwagen et al, 2003). Thus, orientation of Ku80 away from the end at DSBs may sequester Ku80-bound telomerase from a potential substrate (Figure 1-8C).

Inhibition of DNTF is also achieved by the resectioning of DSBs by the exonucleases Exo1 and Sgs1, recruited by the MRN/X complex (Figure 1-8C) (Chung et al, 2010; Lydeard et al, 2010). In budding yeast, loss of both Exo1 and Sgs1 results in

minimal DSB resectioning leading to increased Cdc13 binding, and increased DNTF (Marrero and Symington, 2010; Chung et al, 2010; Lydeard et al, 2010). Thus, resectioning at DSBs, a process that Ku minimizes, inhibits DNTF.

Another mechanism to inhibit telomerase activity at DSBs is through the kinase ATR. ATR is alerted to a resected DSB by RPA-bound ssDNA (Zou and Elledge, 2003). Once activated, ATR phosphorylates two proteins to produce two different outcomes. One protein is Pif1, a 5'-3' DNA helicase and potent telomerase inhibitor (Zhou et al, 2000; Boule et al, 2005). Pif1 physically associates with telomerase *in vivo* and dislodges it from telomeric DNA, thus preventing DNTF (Figure 1-8C) (Eugster et al, 2006; Boule et al, 2005). Another phosphorylation target of ATR is the telomere capping component Cdc13 (Tseng et al, 2006). Phosphorylation of Cdc13 at S306 result in decreased associated of Cdc13 at DSBs, and reduced DNTF (Zhang and Durocher, 2010). Finally, ATR-mediated suppression of telomere components and MRN/X association at DSBs represents a concerted effort to suppress DNTF in yeast. It is not clear if a similar mechanism operates in other organisms. This question will be addressed in chapters II and VI.

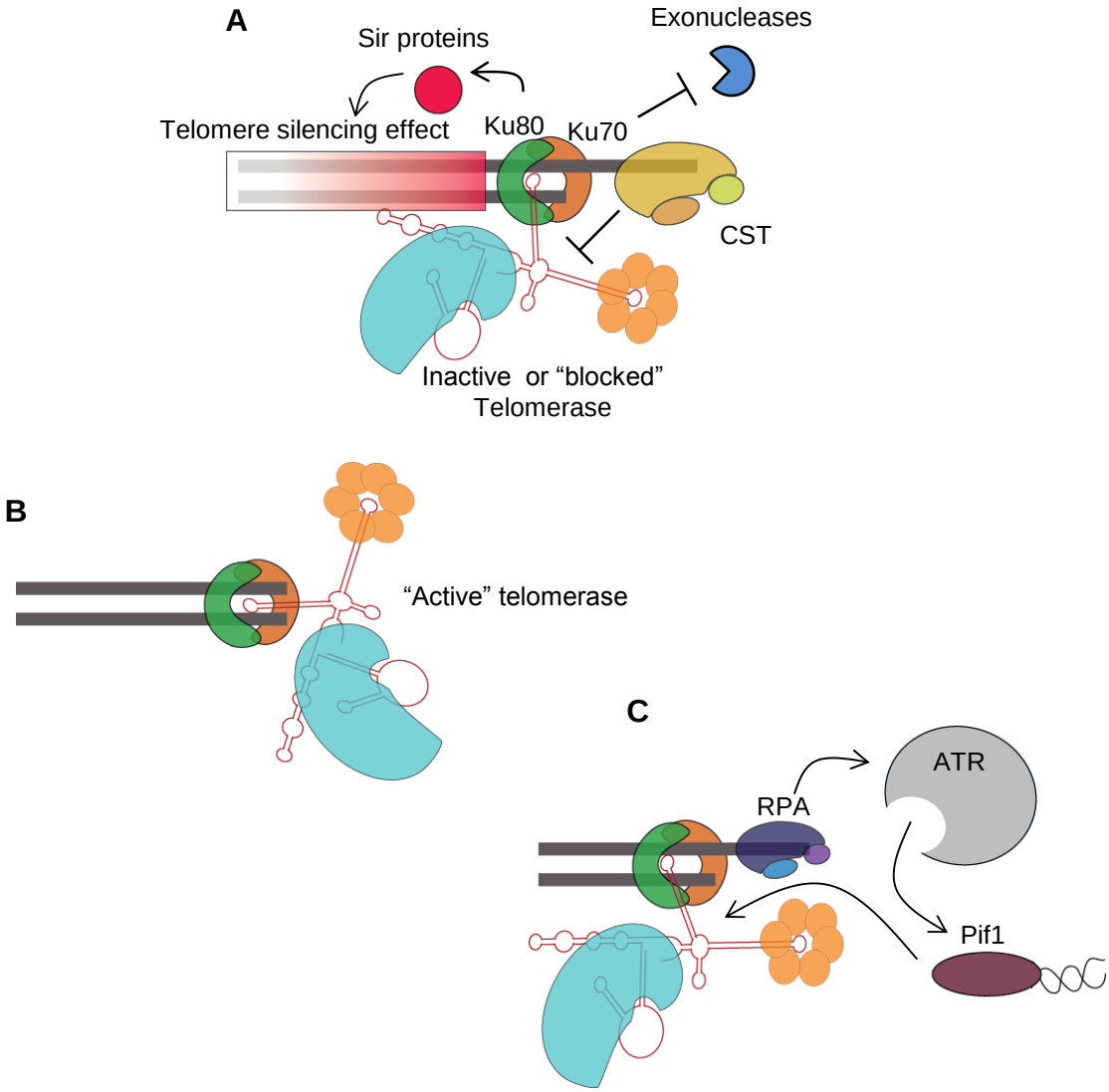


Fig 1-9. A model for Ku activities at telomeres and DSBs in yeast. (**A**) Ku is responsible for normal telomere maintenance. Ku naturally associates with telomeres, where it modulates different activities through its "two faces." Ku binds DNA so that Ku70 faces outward, towards the terminus, whereas Ku80 faces inward, away from the telomere or break. Ku70 is responsible for inhibiting exonucleolytic activity at telomeres. The Ku80 face promotes the telomere silencing effect through interactions with Sir proteins (Pink gradient). This silencing effect represses transcription of genes adjacent to telomeres. In addition, Ku80 is responsible for interactions with TER. This interaction results in telomerase recruitment in the G1 phase of the cell cycle, where it remains inactive, due to lack of Est1. (**B**) *De novo* telomere formation at DSBs. At DSBs that have not been resected, the potential arises for Ku dependent recruitment of telomerase and subsequent DNTF. The mechanism for how this occurs is unclear. (**C**) Inhibition of Ku-mediated DNTF. DNTF at resected DSBs is inhibited by RPA activation of ATR. ATR phosphorylates the DNA helicase PIF1, which inhibits telomerase by forcing it off of the DNA. Resectioning of DNA by itself is also inhibitory, perhaps placing the ss 3' overhang out of reach of Ku-bound telomerase.

One mechanism thought to prevent this interaction from taking place at sites of DNA damage in mammals is the post-translational modification or compartmental sequestration of telomerase. Evidence for this scenario has turned up in human endothelial cells. During high levels of DNA damage, TERT is phosphorylated, leading to inhibition or translocation out of the nucleus and into the mitochondria, where it performs an unknown function (Haendeler et al, 2004; Santos et al, 2006; Kharbanda et al, 2000). In addition, ionizing radiation results in telomerase retention in nucleolar compartments, away from DSBs (Wong et al, 2002).

Arabidopsis as a model to study telomere biology

Arabidopsis thaliana derives its name from the physician who first took note of it, Johannes Thal, in 1577. The idea of using *Arabidopsis thaliana* as a model organism dates back to the 1940s. Friedrich Laibach, the pioneer of Arabidopsis research and of flowering time, recognized the abundance of natural variation found within the species (Somerville and Koornneef, 2002; Koornneef , 2012). *A. thaliana* is found on all continents in the northern hemisphere, and comprises of >1000 different ecotypes (www.tair.org). These ecotypes provide a wealth of information to the community, particularly in the age of genome sequencing (Schneeberger et al, 2011; Cao et al, 2011).

Arabidopsis excels as a model eukaryotic system in that it has a small genome (~130Mb) distributed amongst five chromosomes, short growth period (~6-weeks from seed to seed), and is easily transformable. The Arabidopsis community has developed a wide array of genetic resources, including T-DNA insertion lines and EMS mutant

collections. Arabidopsis is also well-suited as a model system for telomere biology. In the Columbia ecotype, telomeres are relatively short, ranging from 2-5 kilobases, making bulk quantification of telomere length straightforward (Shakirov and Shippen, 2004). Moreover, on eight out of ten chromosome arms there are unique subtelomeric sequences that facilitate the precise measurement of telomere length on a per arm basis (Heacock et al, 2007). This feature makes it possible to use PCR strategies to detect and characterize end-to-end chromosome fusions, which heretofore has been limited to cytological analysis. Arabidopsis is also highly tolerant to genome instability. Mutations that are lethal in other organisms, such as perturbations in telomere capping components and ATM/ATR are not lethal in *A. thaliana* (Song et al, 2008; Surovtseva et al, 2009; Riha et al, 2001; Vespa et al, 2005; Culligan et al, 2004).

TERT and TER in Arabidopsis

The telomerase holoenzyme in Arabidopsis shows remarkable similarity in protein composition to both humans and budding yeast. TERT was identified rapidly following the sequencing of the *A. thaliana* genome (Riha et al, 2001; Arabidopsis Genome Initiative, 2000). AtTERT, at roughly 130kD, is similar in size to hTERT. AtTERT is highly expressed in Arabidopsis cell culture and shoot apical meristems (Oguchi et al, 1999). Telomerase activity is robust in young seedlings, continues to be high in the meristematic regions of the plant, peaking in flowers. Telomerase activity is repressed in vegetative tissue (Fitzgerald et al, 1999; this work). Arabidopsis can survive

in the absence of TERT for up to ten plant generations, but late generation mutants have reduced viability and eventually arrest in a small, vegetative state (Riha et al, 2001).

Arabidopsis thaliana encodes two TER isoforms that can assemble with TERT *in vivo* and reconstitute activity *in vitro* (Cifuentes-Rojas et al, 2011; this work). Depletion of TER1, but not TER2, results in telomere shortening *in vivo*. TER1 and TER2 form distinct RNP complexes *in vivo*, the components of which will be discussed below. TER2 undergoes several unique biogenesis steps *in vivo*. The significance of these events will be a major component of this dissertation, and will be discussed in chapters II-V.

Telomerase accessory factors in Arabidopsis

Several accessory factors have been described for Arabidopsis telomerase. The H/ACA box binding factor, Dyskerin, associates with TER1 and TER2 and is critical for telomerase biogenesis, just as in humans (Kannan et al, 2008). As discussed above, Ku associates with the yeast and human telomerase RNA (Stellwagen et al, 2003; Ting et al, 2005). As described in chapter II, evidence is provided to show that Ku associates with the TER2 RNP in Arabidopsis. In striking contrast to the situation in yeast where loss of Ku results in telomere shortening (Porter et al, 1996; Boulton and Jackson, 1996), loss of Ku in Arabidopsis results in rapid extension of telomeres and increased telomere recombination (Riha and Shippen, 2003; Zellinger et al, 2007).

The best characterized telomerase accessory factor in *A. thaliana* is POT1a. AtPOT1a was identified in a homology search using *S. pombe* POT1 as the search query

(Baumann and Cech, 2001). AtPOT1a retains the requisite secondary structure of a POT1 protein, having two OB-folds and a C-terminal protein-protein interaction domain (Baumann and Cech, 2001; Surovtseva et al, 2007). In striking contrast to its counterparts in yeast and vertebrates, POT1a does not display DNA binding and instead shows high specificity for TER1 (Surovtseva et al, 2007; Cifuentes-Rojas et al, 2011). AtPOT1a is enriched at telomeres during S-phase when telomerase extends telomeres. AtPOT1a also interacts with the telomere capping component CTC1, a functional homolog of Cdc13 (Surovtseva et al, 2007; Beilstein et al, in preparation). Null mutations in AtPOT1a (*pot1a-1* and *pot1a-2*) lead to an EST phenotype, indicating POT1a is required for telomere maintenance (Surovtseva et al, 2007). While mutations in AtPOT1a lead to reduced telomerase activity, TER1 levels are not affected. Thus, POT1a is an essential component of telomerase, necessary for optimal enzymatic activity, but not for RNA biogenesis. Whether POT1a serves as a telomerase recruitment factor similar to *S. cerevisiae* Est1 is still unknown.

In *A. thaliana*, the POT1 gene family has undergone two duplication events, resulting in two full-length POT1 proteins (AtPOT1a and AtPOT1b; Shakirov et al, 2005), and a third, truncated OB-fold encoding gene (POT1c). Over-expression of the N-terminal half of POT1b results in a stochastic decrease in telomere length, chromosomal fusions, and severe morphological defects (Shakirov et al, 2005). Surprisingly, a null mutation within POT1b does not replicate this phenotype. Interestingly, *Physcomitrella patens*, a basal plant species, contains a single ssDNA-binding POT1 protein that performs a canonical role at telomeres (Shakirov et al, 2010). Thus, the potential remains

for an AtPOT1 protein playing a protective role at the chromosome end. The role of AtPOT1b and POT1c in Arabidopsis telomere biology will be discussed in chapters IV and V.

Telomere capping proteins in Arabidopsis

Until recently, none of the factors needed to protect telomeres in *A. thaliana* were known. The identification of STN1 in *S. pombe* was pivotal in revealing essential telomere protection proteins in Arabidopsis (Martin et al, 2007). A BLAST search using the *S. pombe* STN1 sequence as a query revealed the *A. thaliana* STN1 ortholog (Song et al, 2008). STN1 proved to be a conserved telomere capping protein, serving many of the functions of its yeast counterparts. Plants deficient for STN1 showed dramatic telomere loss, massive chromosomal fusions, and increased G-overhang signal (Song et al, 2008). Soon afterwards, while analyzing an EMS mutagenesis line with a mutation within POT1c, a severe chromosome defect was found. The mutation responsible for the telomere phenotype was subsequently mapped, not to POT1c, but to a novel gene, CTC1 (Conserved Telomere maintence Component 1) (Surovtseva et al, 2009). CTC1 mutants showed a profound telomere uncapping phenotype, similar to *stn1*. In addition, it was shown that CTC1 interacts with STN1 *in vitro*. TEN1 is also present in Arabidopsis (Leehy et al, in preparation). Loss-of-function mutants show a similar physiological phenotype to *ctc1* or *stn1* mutants (Hashimura and Ueguchi, 2011). As expected TEN1 interacts with CTC1 and STN1 and is critical for proper telomere capping.

In Arabidopsis, the only other shelterin components, other than the POT1 proteins, are a family of 12 TRF-like (TRFL) proteins, six of which (Class I TRFL proteins) contain a Myb-extension that confers specific ds telomeric DNA binding (Karamysheva et al, 2004). Class I TRFL proteins appear to be good candidates for TRF orthologs in plants. Unfortunately, null mutations in any single one of the class I TRFL genes does not result in any apparent defect in telomere length maintenance, suggesting redundancy amongst this gene family. Strikingly, TPP1, TIN2, and RAP1 sequence homologs cannot be detected in plant genomes. As the existence of CTC1 demonstrates, this does not preclude the presence of functionally homologous shelterin components.

Dissertation overview

The overarching goal of this research was to investigate the mechanism of telomerase regulation. Inhibition of telomerase through the telomerase RNA subunit TER2, and the evolution of this RNA, will be discussed in Chapters II and III. Two new telomerase accessory factors, AtPOT1b and POT1c will be discussed in Chapters IV and V. Finally, the development of Arabidopsis as a platform for studying the mechanisms of DNTF will be discussed in Chapter VI.

In Chapter II, the TER2 telomerase RNP is characterized. TERT binds with a higher affinity to TER2, which brings with it accessory components distinct from the TER1 RNP. Furthermore, the TER2/telomerase complex does not productively engage the chromosome end for telomere maintenance. This finding provides a possible mechanism for how TER2 could negatively regulate telomerase. My contribution to this

story was in finding the physiological circumstances under which TER2 acts. Along with a known marker for activation of a DNA damage response, BRCA1, TER2 levels increase rapidly following addition of genotoxic agents. The induction of TER2 correlates with inhibition of telomerase activity. Importantly, telomerase inhibition by DNA damage is abrogated in a *ter2* null background. Therefore, I hypothesize that TER2 prevents telomerase from acting at sites of DNA damage, allowing for the proper repair of damaged DNA.

Chapter III discusses the evolution of the intervening sequence within TER2. The 1001 Arabidopsis genomes project provides an unprecedented view into the micro-evolution of TER2. I showed that the intervening sequence in TER2 is completely missing from 4 out of the 513 Arabidopsis ecotypes sequenced to date. In 59 other ecotypes the intervening sequence is highly degenerate, suggesting selective pressure against this element. The origin of the TER2 intervening sequence is addressed, as is its conservation and expansion within the Brassicaceae family. I hypothesize that insertion of the IS within an ancestral TER resulted in the duplication and translocation of the TERs in Arabidopsis.

In Chapter IV, I examine the role of AtPOT1b in telomere maintenance and TER2 maturation. I characterize the molecular consequences of a null mutation for POT1b (*pot1b-1*). Contrary to results with transgenic plants over-expressing the N-terminal half of AtPOT1b, we found that *pot1b-1* lines bear minimal telomere defects. However, telomerase activity is increased in this background to a degree seen in the *ter2-1*, indicating POT1b is also a negative regulator of telomerase (Chapter II). POT1b

inhibition of telomerase activity is not dependent on TER2. Preliminary analysis of ter2pot1b double mutants demonstrates dramatic telomere dysfunction similar to those observed in *ctc1* or *stn1* backgrounds. Therefore, I hypothesize that AtPOT1b and TER2 work in concert to protect chromosome ends, but regulate telomerase through different mechanisms.

In Chapter V I establish a function for the newest member of the POT1 family, POT1c. POT1c is limited to *A. thaliana*. POT1c represents a partial gene duplication of POT1a, and yet it functions at telomeres in a manner distinct from POT1a. I show that POT1c is a critical telomere capping component and is required for TER2 stabilization *in vivo*. These findings provide an important functional link between telomerase components and chromosome end protection and highlight the conservation of POT1 function among eukaryotes.

Chapter VI introduces a novel means of investigating DNTF in Arabidopsis. By introducing telomere repeat arrays (TRAs) into tetraploid (4X) Arabidopsis, the baseline *de novo* telomere formation frequency (percent DNTF/population of transformants) was assessed. DNTF was analyzed in different genetic backgrounds to determine what factors regulate DNTF in Arabidopsis. I found that *de novo* telomeres are treated like endogenous telomeres in terms of telomere length regulation. Furthermore, I found that DNTF is dependent on Ku, but not on other NHEJ components. I hypothesize that Ku is necessary for the stabilization and recognition of the TRAs. Additionally, these findings lay the foundation for understanding how DNTF is regulated in multicellular eukaryotes.

Chapter VII contains the overall conclusions from my work and a discussion of future directions for this research. This chapter touches upon testable models for how the AtPOT1 proteins perform their roles at telomeres and within the telomerase RNP. Also included are ideas for how the Arabidopsis DNTF assay can be exploited to address fundamental questions related to telomerase regulation and chromosome end protection.

CHAPTER II

AN ALTERNATIVE TELOMERASE RNA REPRESSES ENZYME ACTIVITY IN RESPONSE TO DNA DAMAGE

Summary

Telomerase replenishes telomere tracts by reiteratively copying its RNA template, TER. *Arabidopsis thaliana* encodes two TER subunits, but only TER1 is required for telomere maintenance. Here we show that TER2 functions as a novel negative regulator of telomerase *in vivo*. Disruption of *TER2* increases enzyme activity, while *TER2* over-expression inhibits telomere synthesis from TER1. Moreover, unlike TER1 RNP, TER2 RNP cannot efficiently extend telomeres. TER2 is spliced and its 3' end cleaved to generate a third TER isoform, $TER2_S$. The catalytic subunit TERT binds TER2>TER1>$TER2_S$ *in vitro*, suggesting that TER2 could act as a competitive inhibitor to sequester TERT in a non-productive complex *in vivo*. Finally, TER2, but not TER1, accumulates in response to genotoxic stress, leading to TER2-dependent telomerase inhibition. These data reveal a complex regulatory network for Arabidopsis telomerase wherein an alternative TER works in concert with the canonical RNA template to regulate enzyme activity and promote genome integrity.

Introduction

Chromosomes must be capped with an ample reserve of telomeric DNA to ensure genome stability. The telomerase reverse transcriptase facilitates homeostasis of telomere tracts using its catalytic subunit TERT to reiteratively copy the internal RNA

template TER, thereby replenishing sequences lost during DNA replication. Telomerase is a highly regulated enzyme; its action is largely confined to, and essential for, self-renewing cell populations. Inappropriate telomerase expression promotes tumorigenesis, while insufficient enzyme activity triggers genome instability and stem cell-related disease (Artandi and DePinho, 2010). Consequently, sophisticated mechanisms evolved to modulate telomerase activity.

Although TERT is a highly conserved constituent of telomerase, TER subunits have diverged dramatically and exhibit little sequence similarity and vastly different sizes that range from 150nt in *Tetrahymena* to >1800nt in yeast (Greider and Blackburn, 1989; Dandjinou et al, 2004; Leonardi et al, 2008; Webb and Zakian, 2008; Gunisova et al, 2009). Phylogenetic and mutational analysis reveal functionally conserved elements within TER, including a single-strand templating domain typically corresponding to one and a half telomeric repeats flanked by a 5' boundary element and a 3' pseudoknot domain (Solymosy et al, 2002; Autexier and Greider, 1995; Chen and Greider; 2004; Gilley and Blackburn, 1999; Seto et al, 2003; Tzfati et al, 2000; Gunisova et al, 2009). Human telomerase activity can be reconstituted with only the TER pseudoknot/template region and the CR4/CR5 trans-activation domain (Mitchell J, 2000; Tesmer et al, 1999). Similarly, yeast and Arabidopsis telomerase activity is supported by a "Mini T" version of TER consisting of ~150 nts (Zappulla et al, 2005; Qiao and Cech, 2008; Cifuentes-Rojas et al, in prep). While TER contributes to telomerase catalysis by promoting processivity (Chen and Greider, 2004; Lai et al, 2003) and fidelity (Prescott and Blackburn, 1997), much of the RNA appears to function as a scaffold for accessory

factors that facilitate RNP biogenesis, trafficking and interactions with the chromosome terminus (Wang et al, 2007; Chang et al, 2007; Zaug et al, 2010; Latrick and Cech, 2010; Robart et al, 2010; Robart and Collins, 2010). Vertebrate TERs associate with the RNP maturation complex dyskerin, but budding yeast TER assembles with Sm-proteins and is processed as a snRNA (Collins K, 2006). Notably, biogenesis of *Schizosaccharomyces pombe* TER involves a novel "slicing" mechanism[19] that requires the sequential binding of SM and Lsm complexes (Tang et al, 2012). Other TER binding factors include the KU70/80 heterodimer, which in *Saccharomyces cerevisiae* acts as a positive regulator of telomerase (Boulton and Jackson, 1998; Stellwagen et al, 2003).

Most of what is known concerning telomerase regulation centers on enzyme activation. In human cells transcriptional regulation of the catalytic subunit TERT is the major point of control, although alternative splicing, post-translational modification and intracellular trafficking of TERT also contribute to enzyme regulation (Cifuentes-Rojas and Shippen, 2012). Increased expression of TER is correlated with enzyme activation in some settings, but evidence that TER plays a significant role in modulating enzyme activity is currently lacking (Yokoi et al, 2003; Angelopoulou et al, 2008).

In conjunction with tight regulation of telomerase activity at natural chromosome ends, the enzyme must also be strictly prohibited from acting at sites of DNA damage to prevent "chromosome healing" by de novo telomere formation (DNTF). DNTF is a perilous endeavor due to loss of flanking DNA. In humans DNTF is associated with several genetic disorders, including α-thalassemia and mental retardation [46,50]. DNTF

can be curtailed by impeding telomerase interaction with a double-strand break (DSB). As part of the DNA damage response (DDR) in yeast, the telomere capping protein Cdc13 is phosphorylated by Mec1 (ATR), thereby preventing its association with a DSB and the subsequent recruitment of telomerase (Pennock et al, 2001; Zhang and Durocher, 2010). Mec1 also stimulates the phosphorylation-dependent activation of Pif1 [48], a helicase that evicts telomerase particles engaged in synthesis by unwinding the TER-DNA hybrid. While similar mechanisms have not been described for multicellular eukaryotes, human TERT is phosphorylated by c-Abl in response to ionizing radiation, a modification that is associated with telomerase inhibition (Kharbanda et al, 2000). Ionizing radiation also triggers the transient sequestration of hTERT in the nucleolus, a response that would temporarily block DNTF.

Plants control telomerase activity in a similar fashion as animals, repressing the enzyme in leaves where cell division is waning and increasing expression in seedlings, flowers and other cells with high proliferation potential (Fitzgeral et al, 1999). As in vertebrates, core components of *Arabidopsis thaliana* telomerase include TERT and dyskerin (Fitzgerald et al, 1999; Kannan et al, 2008). However, *A. thaliana* is unique among model organisms studied to date as it encodes two telomerase RNA subunits, TER1 (748 nt) and TER2 (784 nt) (Cifuentes-Rojas et al, 2011). TER1 and TER2 share a 220 nt highly conserved domain that in TER2 is divided into two segments interrupted by 529 nt intervening sequence. Both TER1 and TER2 assemble with TERT to form an active enzyme *in vitro*, but only TER1 is required for telomere maintenance *in vivo*. A

null mutation in *TER2* does not perturb telomere length homeostasis under standard growth conditions, and hence the function of this RNA has been unclear.

Another key component of the TER1 RNP is POT1a (Protection Of Telomeres), one of three POT1 paralogs in *A. thaliana* (Surovtseva et al, 2007; Shakirov et al, 2005; Chapters IV and V). Vertebrate and yeast POT1 bind the 3' overhang on the chromosome terminus, thereby prohibiting a DDR and the inappropriate enzymatic reactions triggered by it (Baumann and Cech, 2001; Lei et al, 2004). In contrast, Arabidopsis POT1a is a constituent of the telomerase RNP that contacts TER1 and acts in the same genetic pathway as TERT for telomere maintenance (Surovtseva et al, 2007; Cifuentes-Rojas et al, 2011). Over-expression studies suggest that POT1b contributes to genome integrity (Shakirov et al, 2005), and yet like POT1a, POT1b does not bind telomeric DNA *in vitro* (Shakirov et al, 2009). In addition, POT1b does not bind TER1 (Cifuentes-Rojas et al, 2011). Thus, TER2 and POT1b do not promote the canonical role of telomerase in telomere maintenance.

Here we describe a new regulatory pathway for the telomerase RNP wherein TER2 inhibits telomere synthesis by TER1 in response to DNA damage. Specifically, we show that telomerase activity is elevated in the absence of TER2, and repressed when TER2 is over-expressed. We show that TER2 assembles with POT1b into an alternative RNP distinct in protein composition from TER1 RNP. Our data further suggest that TER2 competes with TER1 for TERT *in vivo*, sequestering the catalytic subunit in a non-productive complex. Finally, we demonstrate telomerase activity is repressed in response to genotoxic stress, and this regulation is dependent on TER2. We conclude

that TER2 is novel component of the DDR that modulates telomerase action to promote genome integrity.

Materials and methods

Primer extension

Primer extension was carried out on total RNA extracted from Arabidopsis cell culture. 0.25pmol of 5‘ end labeled oligonucleotide was incubated with total RNA at 95°C for 5 min and allowed to anneal in two sequential 15 min incubations at 72°C, and 60°C, after which extension mix (50mM Tris-HCl pH 8.3, 15mM KCl, 3mM MgCl2, 5% DMSO, 1mM DTT, 1mM dNTPs, 1.5U RNAseOUT® and 200U SuperScript III® reverse transcriptase (Invitrogen) was added to a 30µl final volume. The reactions were incubated at 58°C for 3 h. The enzyme was inactivated at 80°C for 10 min. The RNA was hydrolyzed by incubation at 70°C with 15µl of 1N NaOH for 10 min. The reaction was neutralized and precipitated with 15µl of 1N HCl, 20µl 3M NaOAc pH 5.2, ethanol and glycogen. The products were resolved by denaturing PAGE.

In vitro telomerase reconstitution

A TERT-pET28a plasmid with an N-terminal T7 tag was used for telomerase reconstitution experiments. Reactions were assembled with 100ng of TERT-pET28a plasmid and 0.5pmol or 0.1pmol of gel purified DNA template encoding TER1 or TER2 respectively, driven by a T7 promoter, in a mix containing Rabbit Reticulocyte Lysate (RRL) (Promega), amino acids, RNase inhibitors, and T7 RNA polymerase. Reactions were incubated for 90 min at 30°C. T7 agarose beads (Novagen) were blocked with

buffer W-100 (20mM TrisOAc [pH 7.5], 10% glycerol, 1mM EDTA, 5mM MgCl2, 0.2M NaCl, 1% NP-40, 0.5mM sodium deoxycholate, and 100mM potassium glutamate) containing 0.5mg/ml BSA, 0.5mg/ml lysozyme, 0.05mg/ml glycogen, 1mM DTT and 1μg/ml yeast tRNA. The reconstitution reaction was mixed with the beads to a 200μl final volume and incubated for 2 h at 4°C with rotation. Beads were washed 6X with 800 μl of W-400 buffer (W-100 containing 400 mM potassium glutamate) and 3X with 800μl of TMG buffer (10mM TrisOAc pH 7.5, 1mM MgCl2, and 10% glycerol). After the final wash, beads were resuspended in 30μl of TMG. 2μl of beads were used for TRAP assays as previously described (130, 242).

In vitro binding assays

For co-IP experiments, POT1a, POT1b, KU70 and KU80 cloned in pET28a with a T7 tag were co-expressed with TERs in RRL as described above. After IP RNA was extracted and RT-PCR was performed (see Supplementary Methods). Electrophoretic mobility shift assays were performed with RNA transcribed *in vitro* with T7 RNA polymerase and [α-32P]-CTP labeled TER1 and TER2. Binding reactions contained 3μl of RRL expressed protein, 0.1pmol of 32P labeled TER and 1X binding buffer (25mM Tris-HCl pH 8.0, 10mM Mg(OAc)2, 25mM KCl, 10mM DTT and 5% Glycerol) in a 30μl final volume. 1μM yeast tRNA and 0.5μM RNA (U3AG3)4 were used as nonspecific competitors. After 20 min at 30°C the reaction was loaded onto a 0.8% agarose 0.5X TBE gel and run for 2 h at 70 volts at 4°C. Gels were dried and exposed to phosphorimager screens.

For double-filter binding assays, TERT expressed in RRL was incubated with

decreasing concentrations of pre-folded RNA transcribed *in vitro* using 5-end 32P-labeled RNA as a tracer. Binding reactions contained 0.5µl of recombinant protein, pre-folded TER in binding buffer (50mM Tris-HCl pH 7.5, 200mM potassium glutamate, 0.5mg/ml BSA, 0.5mg/ml tRNA, 1mM MgCl2, 1mM DTT and 0.01% NP-40) in a 25µl final volume. After 30 min at 30°C, the reactions were filtered through nitrocellulose and nylon filters using a dot-blot apparatus (BioRad). The membranes were washed with 600µl washing buffer (50mM Tris-HCl pH 7.5, 200mM potassium glutamate, 1mM MgCl2, 1mM DTT and 10% glycerol), dried, exposed to a phosphor storage screen and scanned after 2h. Equilibrium dissociation constants, Kd, were obtained by non-linear regression of the binding data fitted to a one-site binding model using Graphpad Prism software.

Plant materials and genotyping

Arabidopsis seeds with a T-DNA insertion in TER2 (SAIL_556_A04) were obtained from the Arabidopsis Biological Resource Center (Ohio State University, Columbus, OH). Seeds were cold-treated overnight at 4°C, and then placed in an environmental growth chamber and grown under a 16-h light/8-h dark photoperiod at 23°C. Plants were transformed using the in planta method [88]. For genotyping, DNA was extracted from flowers and PCR was performed with the following sets of primers: #38: 5'-GACGACAACTAA ACCCTACGCTTACA-3' and #45: 5‘-CGATGTTGTTTTTCTGCTTAGGACACA-3‘. To amplify mutant TER2 alleles containing a T-DNA insertion, the T-DNA specific primer was used along with TER 8526-01 fwd: 5‘-GAGACGCAGCGAGCGATAGCCGATAG-3‘ primer.

Template mutation and plant transformation

To generate template mutations in the template region of TER2, a PCR product containing the desired template mutation was generated with the primers TER2RSA fwd (5‘-CACCGACGACAACTAGTACCTACG CTTACA-3‘) and TER2 end reverse (5‘-AATTCTGTGTAGCTATGATCTTGTGGCA-3‘). The mutation was confirmed by sequencing. TER2RSA was cloned into the destination vector pB7WG2 and transformed into plants homozygous for the T-DNA insertion in ter2-1. After transformation the seeds were selected in MS agar containing kanamycin at 50µg/ml.

End-point RT-PCR and quantitative RT-PCR

Total RNA was extracted from 0.5 g floral or other tissue using Tri Reagent (Sigma). cDNAs were synthesized from total RNA using Superscript III reverse transcriptase (Invitrogen). Random pentadecamers were incubated with 2 µg of total RNA in the supplied buffer at 65°C for 5 min. Reverse transcription (RT) was carried out with 100U of Superscript III at the following temperatures 37°C for 20 min, 42°C for 20 min and 55°C for 20 min. Enzyme was inactivated at 80°C for 10 min and RNA was degraded with RNase H (New England Biolabs). 1.5 µl of cDNA was used in PCR. For real-time RT-PCR, 2 µl of the above cDNA was used at a 1:10 dilution in a 20µl reaction containing 10 µl of SyBr green master mix (NEB) and 2 µl of each primer (2µM). PCR was performed for 40 cycles with 30 s at 95°C and 60 s at 60°C. Threshold cycle values (Ct) were calculated using an iCycler iQ thermal cycler (BIO-RAD) and the supplied Optical System Software.

Quantitative RT-PCR data analysis

Amplification efficiencies were calculated for each primer pair in a 5 point titration curve. The slope was calculated from a standard curve where Ct was on the y-axis and log(cDNA dilution factor) on the x-axis. The corresponding real-time PCR efficiency (E) was calculated according to the equation: E=(10-1/slope)-1. To correct for intra-assay and inter-assay variability, each sample was evaluated by tripicate within one run in at least three different experimental runs. The relative expression level (R) was calculated as follows: R = (Etarget)ΔCttarget(control-sample) / (Ereference)ΔCtreference(control-sample) as previously described (329). U6 snRNA and β-actin were used as reference. Normalization to the pre-immune control and to the efficiency for each antibody was used for RNA quantification in the pull-down samples. Primers used for real-time PCR are as follows: TER1 Q4F:5‘-CCCATTTCGTGCCTATCAGACGAC-3‘. TER1 Q4R: 5‘-TCTCCGACGACCATTCTCTCGATAC-3‘; TER2#38: 5‘-GACGA CAACTAAACCCTACGCTTACA-3‘ and TER2#40: 5‘-CAGGATCAATCGGAG AGTTCAATCTC-3‘; TER2S: TER2#38 and TER2S# 193: 5‘-CCCCATCTCCGA CGAGACGAC-3‘; TERT Q3F: 5‘-ACCGTTGCTTCGTTGTACTTCACG-3‘ and TERT Q3R: 5‘-CGACCCGCTTGAGAAGAAACTCC-3‘; U6-1F: 5‘-GTCCCTTCG GGGACATCCGA-3‘ and U6-1R: 5‘-AAAATTTGGACCATTTCTCG A-3‘ β-Actin 2F: 5‘-TCCCTCAGCACATTCCAGCAGAT-3‘ and β-Actin 2R: 5‘-AACGATT CCTGGACCTGCCTCATC-3‘.

TRF analysis, TRAP and Q-TRAP assays

TRF and TRAP assays were performed as previously described (Shakirov et al, 2005; Kannan et al, 2008). TRAP products from mutant TER2RSA telomerase, a specific mutant reverse primer 5‘-CCTAGTACCTAGTACCTAGTACCTA-3‘ was used. Q-TRAP was performed as previously described (Kannan et al, 2008).

Antibodies, immunoprecipitation and western blotting

AtKU70 antibodies were kindly provided by Dr. Karel Riha (Gregor Mendel Institute, Vienna). AtTERT and AtPOT1a antibodies have been previously described (130, 242). The anti-dyskerin polyclonal antibody was raised in rabbits against recombinant full-length AtNAP57 expressed in E. coli. The POT1b is an affinity-purified peptide antibody (Covance). IP efficiency was calculated for each antibody using 35S labeled protein expressed in rabbit reticulocyte lysate. Western blotting was performed with a 1:2000 dilution of anti-KU70, anti-POT1a, anti-POT1b, anti-TERT and anti-Histone H3 antibodies (Upstate). The anti-dyskerin antibody was used at a 1:5000 dilution. Peroxidase-conjugated light chain-specific mouse anti-rabbit secondary antibody (Jackson Immunoresearch) was used at a 1:20,000 dilution. Following IP, RNA was phenol:chloroform extracted from the beads and subjected to RT using superscript III® reverse transcriptase (Invitrogen) and random pentadecamers.

Results

A third TER isoform is generated by splicing and 3' end cleavage of TER2

We unexpectedly uncovered a third isoform of *A. thaliana* TER in experiments designed to examine the expression profile of TER1 and TER2. Primer extension with an oligonucleotide complementary to a region conserved in both TER1 and TER2 (R2) (Figure 2-1A) generated the predicted products as well as a smaller species of ~220 nt (Figure 2-1B). This new RNA was amplified by end-point RT-PCR and was also detected by northern blot analysis (Figure 2-1C), ruling out artifactual PCR amplification or bypass reverse transcription. For reasons discussed below, the new RNA was termed $TER2_S$.

Cloning and sequencing revealed that $TER2_S$ is identical to TER2 with two notable exceptions. First, $TER2_S$ lacks the 529 nt segment in TER2 that interrupts the two highly conserved regions shared with TER1 (R1 and R2) (Figure 2-1D). In $TER2_S$ R1 and R2 are precisely joined to create a contiguous 220 nt stretch with 85% identity to the corresponding region in TER1. The 11 nt telomere template sequence is retained in $TER2_S$. BLAST searches failed to identify a locus in the *A. thaliana* genome that could encode $TER2_S$, indicating that this RNA is a processed form of TER2.

Second, $TER2_S$ lacks 36 nt from the 3' terminus of TER2 (Figure 2-1D). The missing segment corresponds to a non-conserved region just downstream of R2. Characteristic landmarks of consensus mRNA splicing (branch point, 5' and 3' splice sites) are not detected within the 529 nt intervening sequence or in the sequence flanking the 3' cleavage site in TER2 (Figure 2-1E), suggesting that $TER2_S$ is not generated by conventional mRNA splicing. Since the three TER isoforms were isolated from Arabidopsis cell culture, we assessed the steady state levels of these RNAs during plant development. Quantitative RT-PCR showed that all three RNA isoforms are expressed in Arabidopsis plants. The steady state levels of TER1 and TER2s are consistently higher than TER2, although the abundance of all three RNAs declines in non-reproductive tissues (Figure 2-1E).

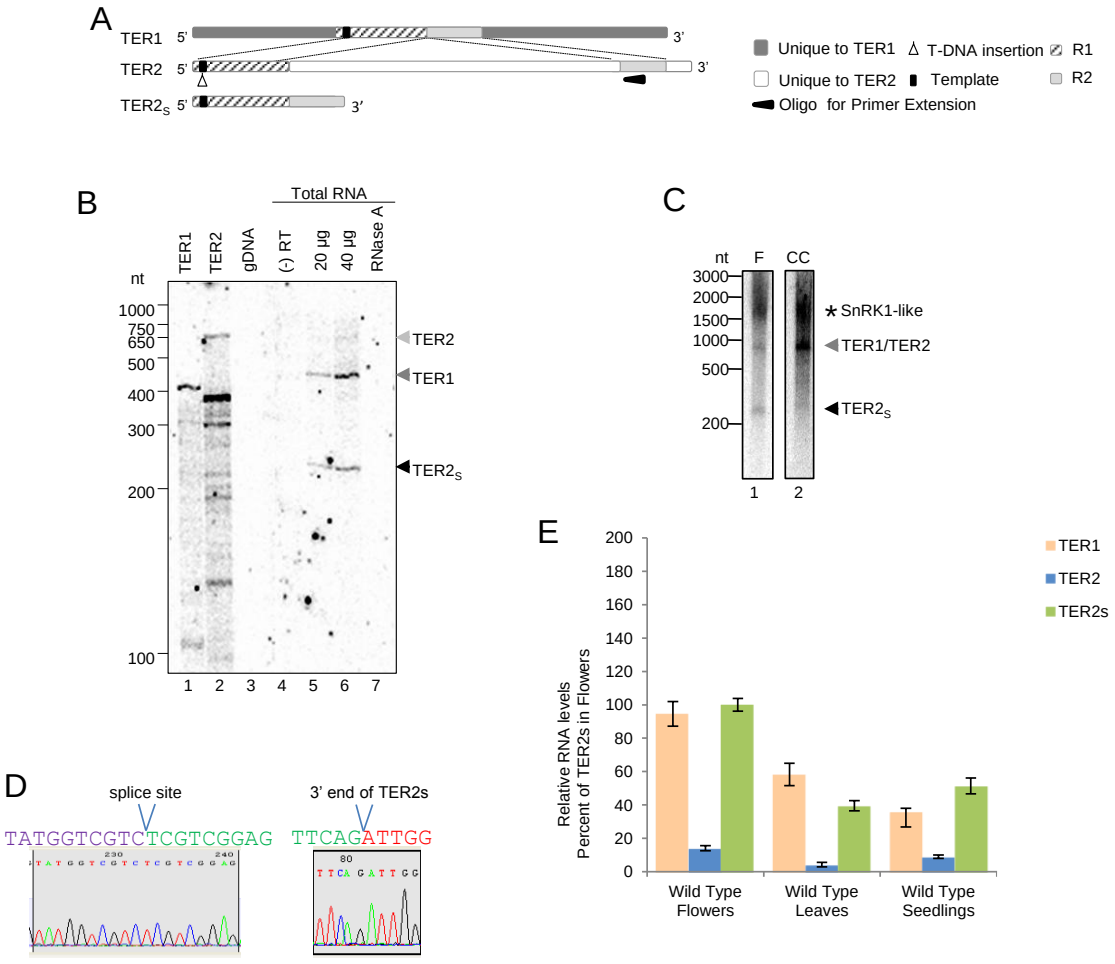


Fig 2-1. *A. thaliana* contains three TER isoforms. (**A**) Primer extension of total RNA from cell using a primer complementary to a region in R2 conserved in TER1 and TER2. Lane 1, *in vitro* transcribed TER1. Lane 2, *in vitro* transcribed TER2. Lane 3, 40 µg genomic DNA (gDNA). Lane 4, control with 40 µg of total RNA in the absence of reverse transcriptase. Lane 5, 20 µg of total RNA. Lane 6, 40 µg of total RNA. Lane 7, 40 µg of total RNA pre-treated with RNase A. (**B**) RT-PCR results with total RNA from cell culture. cDNA was generated using random pentadecamers. Odd lanes correspond to minus reverse transcriptase controls. Reactions with primers to detect specifically TER2 or $TER2_S$ are shown in lanes 1-4 and with primers to detect both TER2 and $TER2_S$ (lanes 7-8). U6 snRNA was amplified as a control (lanes 5-6). (**C**) Northern blot from wild type flowers and cell culture total RNA (20µg, 40µg and 60µg) using a radiolabeled probe complementary to the conserved region in TER2. Black arrow head, $TER2_S$. Grey arrow head, TER1/TER2. Asterisk, SnRK1-like mRNA is also detected because the TER conserved regions are embedded in this predicted coding gene. (**D**) Representative sequencing of TER2s RNAs isolated from cell culture. (**E**) Normalized TER levels. Levels are relative to TER2s in flowers. Error bars represent S.D. n>5 for all data points.

TER2 assembles into an active enzyme, but it cannot efficiently incorporate telomere repeats onto chromosome ends

We previously showed that TER2 assembles with TERT into an active enzyme *in vitro*. Interestingly, however, in the presence of TER2, reconstitution of a catalytically active enzyme is optimal with extremely low RNA concentrations (.001 pmol) compared to 10 pmol for TER1 (data not shown). These findings argue that biochemical properties of TER1 and TER2, and their interaction with TERT *in vitro* are distinct. Consistent with this finding, analysis of a TER2 null mutant, *ter2-1*, revealed that this RNA does not make a significant contribution to telomere maintenance *in vivo*.

Therefore, we asked if TER2 assembles into an enzymatically active RNP *in vivo* by mutating the templating domain from 5'-CUAAACCCUA-3' to 5'-CUAGUACCUA-3' ($TER2_{RSA}$) to differentiate telomere repeat synthesis from TER2 and TER1. The templating function of $TER2_{RSA}$ was intact as *in vitro* reconstitution reactions with TERT yielded the expected RSA-type repeats (Figure 2-2A). $TER2_{RSA}$ was then placed under the control of the powerful Cauliflower Mosaic Virus (CaMV) 35S promoter and transformed into plants null for TER2, *ter2-1*. As expected, TER1 levels were wild type in these lines, while 35S::$TER2_{RSA}$ was over-expressed. Unlike wild type seedlings where $TER2_S$ is much more abundant than TER2, $TER2_{RSA}$ was enriched over $TER2_{S-RSA}$, and thus became the dominant TER2 isoform (Figure 2-2B). TRAP assays were performed on nuclear extracts from the transformants using primers designed to detect RSA-type repeats. Robust telomerase activity was detected, demonstrating that $TER2_{RSA}$ assembled into an enzymatically active RNP *in vivo.*

We used Primer Extension Telomere Repeat Amplification (PETRA) to test if $TER2_{RSA}$ directed telomere repeat incorporation onto chromosome ends. PETRA reactions with $TER2_{RSA}$ transformants yielded the expected products with wild type PETRA-T primer (Figure 2-2C). In a second set of controls, the PETRA-T_{RSA} primer failed to generate products for wild type plants or plants expressing a mutant form of TER1 ($TER1_{CC}$) (data not shown). Very faint products were detected in reactions with PETRA-T_{RSA} for one of the two $TER2_{RSA}$ mutant lines (Figure 2-2C). Sequence analysis revealed a single mutant repeat in 14% of the clones (7/50). In one example, the mutant repeat was flanked by the sequence TATA, which is not encoded by the template. In another clone, the mutant repeat was adjacent to a partial mutant repeat. We suspect that the low level of $TER2_{RSA}$-directed repeats incorporated reflects TER2 over-expression and the resultant inhibition of TER1 (see below). We previously showed that TER1 is a highly efficient telomere template *in vivo* (Cifuentes-Rojas et al, 2011). In striking contrast to the results with $TER2_{RSA}$, more than half of the telomere tracts cloned from plants expressing a mutant TER1 ($TER1_{CC}$) carried mutant telomere repeats, even though $TER1_{CC}$ was competing with endogenous wild type TER1 in this earlier experiment (Cifuentes-Rojas, 2011).

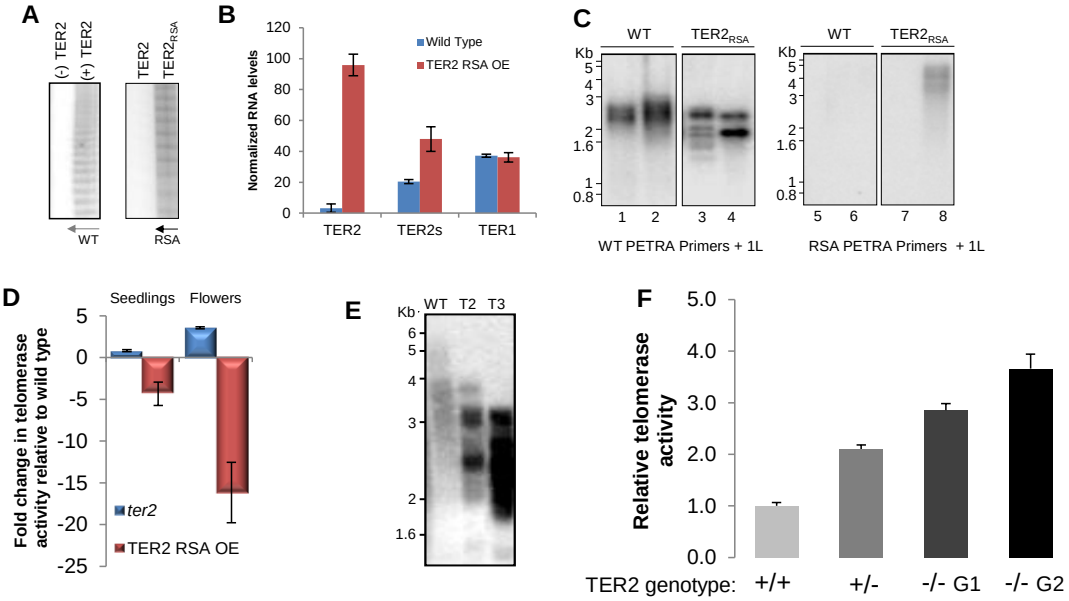


Figure 2-2. **TER2 assembles into an active enzyme *in vivo, but cannot maintain telomere repeats on chromosome ends*** (A) TRAP from $TER2_{RSA}$ *in vitro* reconstituted complexes. (B) Quantitative RT-PCR results for TER levels in $TER2_{RSA}$ transformants (C) TRAP results for 35S::$TER2_{RSA}$ transformants. Grey arrow, reverse primer complementary to WT telomere repeats. Black arrow, reverse primer complementary to $TER2_{RSA}$ mutant repeats. (C) Results of amplification are shown with PETRA T and PETRA-A to detect WT telomeric DNA (left), and with PETRA-T_{RSA} and PETRA-A primer to detect $TER2_{RSA}$ products (right). (D) Q-TRAP results for 35S::$TER2_{RSA}$ transformants. TRAP was performed using a reverse primer complementary to the wild type repeat. (E) TRF analysis of second (T2) and third (T3) generation TER2RSA mutants. (F) Q-TRAP results for WT, *ter2-1*$^{+/-}$, and G1 and G2 *ter2-1* homozygous mutants. Values were normalized to telomerase activity in WT plants. TRF analysis of second (T2) and third (T3) generation TER2RSA mutants. (F) Q-TRAP results for WT, ter2-1+/- and ter2-1-/- plants. Values were normalized to telomerase activity in WT plants.

TER2 negatively regulates TER1-directed telomerase activity in vivo

Q-TRAP assays performed with plants expressing $TER2_{RSA}$ unexpectedly revealed that endogenous telomerase activity was reduced by ~11-fold compared to untransformed controls (Figure 2-2D). Because 35S::$TER2_{RSA}$ transformants harbor a null mutation at the *TER2* locus, telomere repeat synthesis must be suppressed from the TER1 RNP. Although no telomere length change was observed in first generation (T1) transformants (data not shown), terminal restriction fragment analysis (TRF) revealed markedly shorter and less heterogeneous telomere tracts in the second (T2) and third (T3) generations (Figure 2-2E), consistent with reduced telomerase activity *in vivo* (Kannan et al, 2008; Riha et al, 2001). The decrease in telomerase activity correlated with increased levels of $TER2_{RSA}$ and not $TER2_{S\text{-}RSA}$ (Figure 2-2B), arguing that TER2, and not its processed form, $TER2_S$, negatively regulates telomerase activity.

If TER2 is a negative regulator of telomerase, enzyme activity should be elevated in plants null for *TER2*. As predicted, telomerase activity increased 2.1 fold in first generation (G1) *ter2-1* heterozygotes, 2.8-fold in G1 *ter2-1* homozygotes and 3.7-fold increase in second generation (G2) *ter2-1* homozygotes (Figure 2-2F). Telomerase activity was confined to organs where the enzyme is normally expressed; enzyme levels remained low or undetectable in leaves (data not shown). Hence, loss of TER2 did not alter the developmental regulation of TER1 RNP activity. Consistent with previous results showing that over-expression of *A. thaliana* telomerase does not increase telomere length, TRF analysis revealed no change in telomere length for *ter2-1* mutants

(data not shown). Therefore, we conclude that TER2 inhibits the enzymatic activity of TER1 RNP *in vivo*.

TERT differentially binds the three TER isoforms

To investigate how TER2 negatively regulates TER1 RNP, we examined TER2-protein interactions, beginning with TERT. TERT was expressed in rabbit reticulocyte lysate (RRL) and a double-filter binding assay was performed with decreasing concentrations of *in vitro* transcribed TER1, TER2 or $TER2_S$ to calculate Kd values. The Kd for TERT-TER1 and TERT-TER2 was 204.1 ± 11.3 nM and 22.6 ± 2.8 nM, respectively (Fig 2-3A), indicating that the affinity of TERT for TER2 is approximately 8-fold higher than for TER1. The Kd value for TERT-$TER2_S$ was significantly higher than for TER1 at approximately 1.5 µM (Fig 2-3A). Thus, TERT displays a binding preference for TER2>TER1>$TER2_S$ *in vitro*.

Immunoprecipitation (IP) using anti-TERT antibody followed by end-point and Q- RT-PCR were performed with five day-old *A. thaliana* cell culture extracts to study TERT-TER interactions *in vivo*. None of the TER isoforms were detected in Histone H3 or pre-immune IP control reactions (Figure 2-3B). The TERT IP contained both TER1 and TER2, but only trace amounts of $TER2_S$, consistent with in the vitro binding data. Approximately two-fold more TER1 was precipitated relative to TER2 (Fig2-3B), although TER1 is approximately 19-fold more abundant than TER2 in this cell culture extract (Figure 2-3B). The preferential enrichment of TER2 over TER1 in the TERT IP suggests that TER2 could outcompete TER1 for TERT *in vivo*.

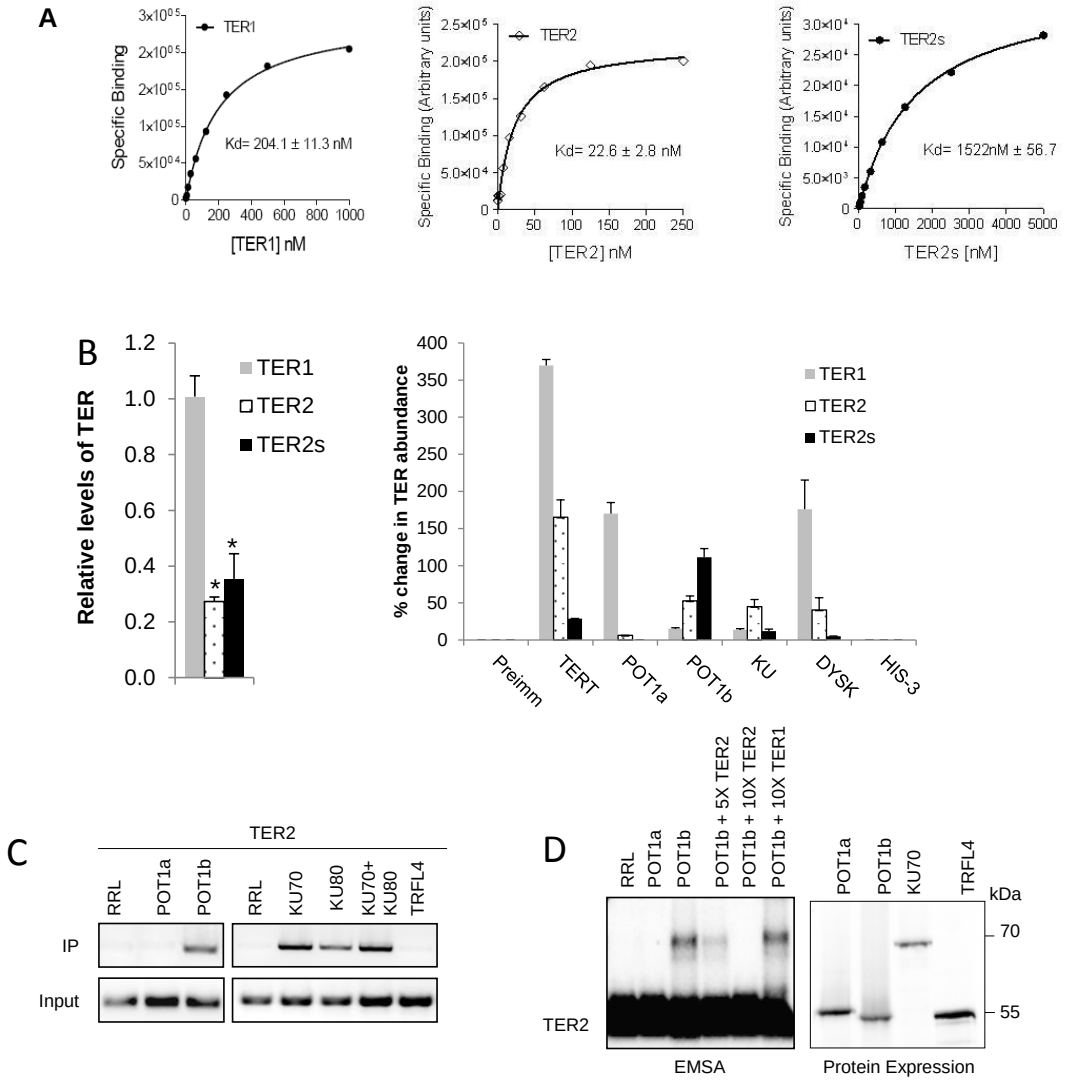


Fig 2-3. The three TERs form distinct telomerase RNPs. (A) The affinity of the TERT-TER1 and TERT-TER2 interaction was determined by double filter binding. TER was added in decreasing concentrations to a reaction containing a constant amount of TERT. Binding isotherm of TERT-TER1 (left), TERT-TER2 (middle), and TERT-TER2s (right) are shown. (**B**) Identification of *in vivo* TER interacting partners. Relative levels of TERs in cell culture as input control (left) were compared to immunoprecipitated RNAs (right). IP was carried out with *Arabidopsis* cell culture extracts using the indicated antibodies. Pre-immune serum, anti-histone H3 and anti-enolase antibodies were used as controls. RNA levels were normalized to the primer efficiency, the levels of U6 snRNA, the pre-immune control and the antibody efficiency. (**C**) In vitro characterization of TER2 interacting partners. T7-tagged proteins were co-expressed with TER2 in RRL. RT-PCR was carried out after immunoprecipitation (IP). TRFL4, a ds telomeric DNA binding protein, was used as control. (**D**) Results from electrophoretic mobility shift assays. Labeled RNAs are indicated. Protein expression is indicated to the right.

TER1 and TER2 assemble with different accessory proteins in vivo

Given the marked differences in TERT affinity for the TER isoforms, we postulated that different proteins assemble with these RNAs *in vivo*. A candidate approach was used to identify proteins that bind TER2 and $TER2_S$. We failed to detect *in vitro* binding of Ku with TER1, however *in vitro* co-IP assays using T7-tagged Ku70/Ku80 revealed an interaction with TER2 (Figure 2-3C). A second potential binding partner for TER2 is POT1b, a paralog of the TER1-binding protein, POT1a (Shakirov et al, 2005). POT1b displayed the opposite affinity of POT1a, binding TER2, but not TER1 (Figure 2-3C and D). Gel shift analysis confirmed the interaction between TER2 and POT1b and demonstrated its specificity (Figure 2-3D). *In vitro* filter binding and gel shift analysis also indicated that POT1b binds $TER2_S$ (data not shown).
IP was conducted to evaluate TER-protein interactions *in vivo* using antibodies directed at POT1b, Ku70 and dyskerin (Cifutentes-Rojas, 2011) Quantitative RT-PCR was performed and the data were normalized to account for differences in IP efficiencies and primer usage. TER2 and $TER2_S$, but not TER1 (Cifuentes-Rojas et al, 2011), were detected in the POT1b IP, consistent with the *in vitro* binding data (Figure 2-3B). Although $TER2_S$ is approximately four-fold less abundant than TER2 in the cell culture extract (Figure 2-3B), $TER2_S$ was enriched in the POT1b IP relative to TER2. These results argue that POT1b associates with both TER2 and $TER2_S$ *in vivo*, with a preference for $TER2_S$. As expected, neither TER2 nor $TER2_S$ were detected in an IP reaction with POT1a antibody (Figure 2-3B).

The Ku70 antibody precipitated TER2, but only a trace amount of $TER2_S$ and TER1. In addition, TER1 and TER2, but not $TER2_S$, were found in the dyskerin IP (Figure 2-3B), a finding consistent with the elimination of a putative H/ACA box during 3' end cleavage of TER2. Altogether, the IP data indicate that *in vivo* TER2 and $TER2_S$ assemble into RNP complexes distinct from TER1 RNP. Specifically, TER2 associates with TERT, KU, POT1b and dyskerin, while $TER2_S$ accumulates in a subcomplex containing POT1b.

DNA damage-induced repression of telomerase activity correlates with increased abundance of TER2

Since a null mutation in TER2 does not affect telomere length homeostasis under standard growth conditions, we explored the possibility that TER2 is a regulatory molecule that modulates telomerase activity in response to genotoxic stress. We recently discovered that telomerase activity is inhibited in *A. thaliana* seedlings treated with the radiomimetic drug zeocin (Boltz et al, 2012). To further investigate how telomerase activity levels are affected by DNA damage, seven day-old seedlings were transferred to liquid culture containing 20 uM zeocin and Q-TRAP was performed at different time intervals beginning 30 minutes after transfer to drug. We verified that this regime elicited DDR by monitoring induction of BRCA1, previously shown to transcriptionally up-regulated in response to double-strand breaks (Boltz et al, 2012). As expected, BRCA1 transcripts were elevated within 30 minutes in the drug, and peaked at 2 hours (Figure 2-4B). Further evidence that a DDR was activated came with propidium iodide staining of the root apical meristem. Stem cell death was clearly evident within 24 hours

of drug treatment (Figure 2-4D). We found that within 30 minutes of zeocin treatment, a statistically significant decrease in telomerase activity was observed, with activity levels declining by ~50% relative to untreated samples (Figure 2-4A). Telomerase activity was decreased to the same degree in all of the time points tested, out to 24 hours, indicating that the response of telomerase to genotoxic stress is sustained.

We next asked if telomerase inhibition correlated with changes in the steady state level of TER using Q-RT PCR (Figure 2-4C). TER1 levels were largely unaffected by zeocin treatment. Similarly, no change in TERT mRNA was observed (Data not shown). In contrast, TER2 increased more than 5-fold after 1 hour of zeocin treatment, and became the predominant isoform of TER. Although $TER2_S$ levels rose slightly after 30 minutes in zeocin, no further accumulation of this RNA occurred with longer treatment. Thus, among the three TER isoforms, TER2 appears to be uniquely responsive to DNA damage.

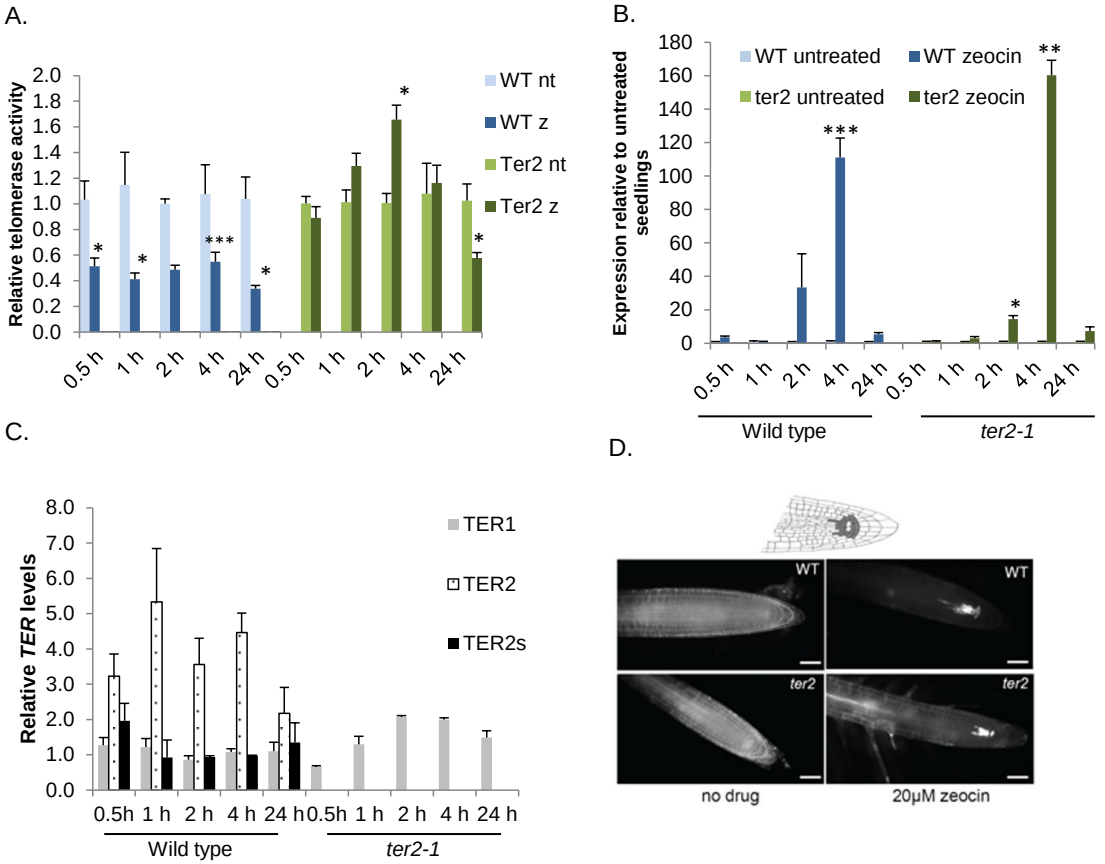


Fig 2-4. A TER2-dependent inhibition of telomerase following DNA damage. (**A**) Telomerase activity is repressed by DNA damage in wild type, but not *ter2-1*, backgrounds. Telomerase activity was determined by the QTRAP assay on seedlings treated for the indicated time points. Nt=not treated, z= with zeocin. Significance calculated relative to untreated at each time point (* = p< 0.05; ** = p<0.005;***=p<0.0005) (**B**) A DNA damage response is activated in the *ter2-1* background. QRT-PCR on BRCA1 RNA levels in treated and untreated seedlings in WT and *ter2-1* backgrounds. (**C**) TER2 RNA levels increase in response to zeocin treatment. Shown is the fold change in RNA levels of zeocin treated seedlings relative to untreated. Genetic backgrounds are shown below the time points. (**D**) Propidium Iodide staining of root tips confirms cell death following zeocin treatment. Root tips of treated seedlings were incubated with propidium iodide, a chemical marker for dead plant cells.

DNA damage-induced repression of telomerase activity is dependent on TER2

Given the correlation between TER2 abundance and telomerase repression, we asked if DNA damage-induced telomerase inhibition is dependent upon TER2. To explore this possibility, we first measured the response of 7 day-old *ter2-1* seedlings to genotoxic stress by monitoring DDR transcript levels and the viability of the root apical meristem at different time points following zeocin treatment. Plants null for *TER2* showed a similar transcriptional to zeocin as wild type plants (Figure 2-4B)

In contrast to wild type, *ter2* seedlings did not display a reduction in telomerase activity following DNA damage. Q-TRAP showed no statistical difference in enzyme activity in *ter2* mutants treated with zeocin for 30 minutes or 1 hour relative to untreated *ter2* controls (Figure 2-4A). In contrast, after 2 hours in zeocin, telomerase activity was significantly elevated in *ter2* seedlings compared to untreated controls. By 24 hours in zeocin, telomerase activity levels were reduced in both *ter2* and wild type seedlings, which at this point may reflect massive stem cell death associated with prolonged genotoxic stress. Altogether, these data argue that TER2 functions as a component of the DDR responsible for modulating telomerase activity in response to genotoxic stress.

Discussion

Gene duplication is a major driving force for genomic diversity. Within the realm of telomere biology, core components of the vertebrate shelterin complex, TRF1/TRF2 and POT1a/POT1b, exemplify how gene duplication and neo-functionalization shape the interactions and regulation of chromosome ends (Price et al, 2010). Critical components

of telomerase have also been subjected to duplication and diversification. The ciliated protozoan *Euplotes crassus* harbors three divergent TERT genes that are expressed at different life cycle stages, prompting a profound switch in telomerase behavior (Karamysheva et al, 2003). Specifically, RNPs containing the TERT-1 and TERT-3 isoforms accumulate during vegetative growth when telomerase engages telomeric DNA to maintain telomeres. In contrast, the TERT-2 isoform is present only during the sexual stage of the life cycle when programmed DNTF occurs on thousands of nascent mini-chromosomes. Remarkably, the gene encoding TERT-2 is degraded at the end of macronuclear development, presumably to prevent DNTF during subsequent vegetative growth. Here we describe another instance of neo-functionalization of a core telomerase subunit in which TER duplication is linked to a novel regulatory pathway that restrains telomerase activity in response to DNA damage.

TER2 processing gives rise to a third TER isoform

The three TER isoforms uncovered in *A. thaliana* are unprecedented; all other organisms studied to date harbor a single TER gene (Dandijinou et al, 2004; Leonardi et al, 2008; Chen and Greider, 2004). TER1 and TER2 are encoded by separate genetic loci (Cifuentes-Rojas et al, 2011), while $TER2_S$ is produced via splicing and 3' end cleavage of TER2. Although a bona-fide splicing reaction has not been described for other TER moieties, cleavage of the 3' end of *Saccharomyces pombe* TER1 is required for telomerase function (Box et al, 2008). SpTER1 transcripts bearing a poly(A) tail can be detected in fission yeast (Leonardi et al, 2008), but the 3' terminus of the RNA

associated with active telomerase is formed by "slicing", a novel mechanism in which the spliceosome carries out only the first transesterification reaction (Box et al, 2008).

Unlike *S. pombe* TER, *A. thaliana* TER2 lacks canonical mRNA splicing signals, arguing that it is subjected to a different set of unconventional RNA processing reactions. Supporting this conclusion, over-expression of TER2 does not lead to a concomitant increase in $TER2_S$, implying that factors necessary for TER2 processing, unlike the conventional spliceosome machinery, are limiting *in vivo*. Further analysis of the RNA processing events that generate TER2S, including exploration of auto-catalysis, clearly warrant further investigation.

TER2 interactions within the telomerase RNP

Our data argue that TER2 controls telomerase RNP function at multiple levels. First, TER2 may act as a competitive inhibitor for the telomerase catalytic subunit. TERT has a higher affinity for TER2 than for TER1 or $TER2_S$ *in vitro*, and preferentially assembles with TER2 *in vivo*. Furthermore, over-expression of TER2 results in inhibition of TER1 mediated telomere-repeat incorporation, consistent with a reduction in the concentration of functional TER1 RNP

Second, several lines of evidence indicate that the TER2 RNP does not productively engage the chromosome terminus, and thus could sequester TERT in an inactive complex. Although reconstitution data indicated that TER2 and TERT can form an enzymatically active particle *in vitro,* the TERT-TER2 interaction is clearly biochemically distinct from TERT-TER1. Enzymatically active telomerase particles can be detected only with very low concentrations of TER2 (1-10 pmol) versus an optimum

for TER1 of 100 pmol. While the molecular explanation for this unusual concentration-dependence is unknown, analysis of the TER2 processing reaction provides some clues.

Preliminary data indicate that TER2 processing (splicing and 3' end cleavage to generate $TER2_S$) proceeds auto-catalytically *in vitro* (Cifuentes-Rojas, A. Hernandez, H. Zu and D. Shippen, unpublished data). Notably, the $TER2_S$ product is generated only at very low concentrations of TER2 RNA (.0001 pmol). Thus, the enzymatically activity generated with TER2 RNA may instead reflect the formation of TERT-$TER2_S$ particles, not TERT-TER2. Similarly, although we could observed synthesis of RSA-type repeats in protein extracts from 35S::TER2 transformants, it is not known whether these products were synthesized from TER2 or TER2s. What is clear, however, is that over-expression does not drive efficient incorporation of telomere repeats onto chromosome ends from the TER2 template, arguing that TER2 and its processed form $TER2_S$ are defective for telomere maintenance *in vivo*. We postulate that for TER2, the presence of a large 529 nt intervening sequence may preclude it from forming an active catalytic core with TERT. In the case of TER2s, this minimal TER may simply lack the binding site for a bridging factor akin to Est1 from budding yeast (Evans and Lundblad, 2002) to properly position the enzyme at the telomere. Because we detect a low level of TER2/TER2s-mediated telomere repeat incorporation, it is possible that these RNP complexes engage the telomere, but in a non-productive mode (see below).

TER2 and the DDR

If TER2 cannot function in telomere length regulation, why did Arabidopsis retain this RNA and assemble it into an elaborate alternative RNP with protein subunits

distinct from its canonical telomerase enzyme? We hypothesize that the role of TER2 is to facilitate error-free DNA repair by down-regulating telomerase activity in response to DNA damage. Telomerase activity is repressed upon the introduction of double-strand breaks or replicative stress, which can ultimately lead to chromosome breaks. In this setting telomerase inhibition would promote genome integrity by reducing the probability of DNTF at sites of DNA damage. Conversely, telomerase activity is not diminished in the face of telomere deprotection, even though a potent DDR is activated. Indeed, robust telomerase activity would be advantageous as it could delay the onset of catastrophic telomere failure. The fate of telomerase may therefore hinge on the chromosomal context in which DDR is triggered.

The steady state level of TER2, but not TER1, TER2 or other mRNAs encoding telomerase-associated proteins, rapidly increases in response to DNA damage. TER2 peaks within one hour of zeocin treatment, concomitant with transcriptional activation of the major DDR marker BRCA1. While the underlying mechanism for the increase in TER2 is unknown, it becomes the major TER isoform. Intriguingly, TER2s does not increase in parallel to TER2, arguing that TER2 processing or stability is affected by genotoxic stress. Most importantly, DDR-induced telomerase inhibition is abrogated in plants lacking TER2. How does TER2 trigger DDR-induced telomerase inhibition? The rapid kinetics of telomerase repression are inconsistent with an exchange of TER1 for TER2 in RNP complexes, although this possibility has not been yet tested. The data are more consistent with the intriguing possibility that TER2 acts as an intermediary within the DDR.

Kedde et al reported a correlation between human TER and the DNA damage response (Kedde et al, 2006). The authors show that hTR, but not hTERT is upregulated in response to UV radiation. Moreover, hTR abundance is inversely correlated with the activity of the checkpoint kinase ATR. Unlike the very rapid rise in Arabidopsis TER2 levels, however, induction of hTR takes place over the course of several hours, suggesting that increased hTR correlates with resumption of the cell cycle following DNA repair, rather than an early event in DDR.

There is mounting evidence in plants and in vertebrates that non-coding RNAs play pivotal roles in genome defense (Hung et al. 2011). The expression, processing and maturation of a variety of miRNAs is altered in response to double-strand breaks in humans (Hu and Gatti, 2011). Some miRNAs suppress H2AX (Lal et al, 2009), ATM (Hu et al, 2010a) and p53 (Hu et al, 2010b; Le et al, 2009), while other others modulate cell cycle progression as part of the DDR (Ivanovska et al, 2008; Wang et al, 2009; Hung et al, 2011). Also induced by DNA damage are long non-coding RNAs such as PANDA, which is implicated in the control of apoptotic gene expression (Hung et al, 2011). Recent exciting studies in Arabidopsis demonstrate a critical role for siRNAs in the recognition and repair of double-strand breaks (Wei et al, 2012).

With its crucial role in stabilizing chromosome termini, telomerase is a likely target for regulation by a DDR-induced non-coding RNA. TER2 falls into the category of long intergenic non-coding RNAs (lincRNA). Typically, lincRNAs act in trans to control gene expression through the recruitment of chromatin remodeling complexes (reviewed in Mercer et al, 2009). Therefore, it is conceivable that TER2 acts as recruiter for telomerase RNP components to the telomere. For example, the interaction of TER2 with POT1b, a molecule previously implicated in promoting telomere integrity (Shakirov et al, 2005), may belie a role for TER2 RNP in stabilizing natural chromosome ends. Similarly, TER2 may promote Ku-mediated inhibition of telomere elongation (Riha et al, 2003). Finally, Ku could bring the TER2 RNP to sites of DNA damage, thereby blocking access of TER1 RNP and reducing the probability of inappropriate telomere formation.

CHAPTER III

EVOLUTION OF THE TELOMERASE RNA INTRON IN ARABIDOPSIS

Summary

Telomeres consist of repetitive DNA-protein elements that are found at the ends of linear chromosomes. Telomeres are terminated by a ssDNA 3' G-rich overhang termed the G-overhang. Telomere homeostasis is a carefully regulated process involving the engagement and extension of this G-overhang by a ribonucleoprotein comple called telomerase. At its core, telomerase consists of a reverse transcriptase, TERT, and an RNA subunit, TER. TERs are defined by their telomere template region, which TERT uses to encode for the reverse transcription of telomeres. Aside from their template region, TERs are long, non-coding RNAs that act as scaffolds for the recruitment of telomerase accessory factors. In Arabidopsis, a duplication event has resulted in two TERs, TER1 and TER2. TER1 and TER2 share a ~219nt region of high similarity that in TER2 is disrupted by a 529nt intervening sequence (IS). Intriguingly, this IS is removed *in vivo* and *in vitro*, resulting in a TER2 isoform termed TER2s. All three TER isoforms can encode for telomeres *in vitro*, but only TER1 is required for telomere homeostasis *in vivo*. Here, using genome data acquired from the 1,001 Arabidopsis genomes project, I determine that the IS within TER2 shares characteristics with the Class II DNA transposons. Unlike the rest of TER2, the IS is experiencing dramatic varation between the different Arabidopsis ecotypes, with three ecotypes lacking the IS completely. Thus,

telomerase inhibition by TER2 may not be the ancestral function, but arose from the incidental insertion of a transposable element.

Introduction

Telomeres are the tandem, G-rich repeats found at the ends of most eukaryotic chromosomes. Failure to maintain telomeres activates a cellular senescence program, or a case of mistaken identity, where telomeres are perceived as double-strand breaks, and subjected to inappropriate DNA repair activities. The telomerase reverse transcriptase continually synthesizes telomeric DNA in stem and germline cells to avert cellular senescence. Telomerase is a ribonucleoprotein complex minimally consisting of a catalytic subunit TERT, and a template bearing RNA subunit, TER. *In vivo*, these two components associate with numerous accessory factors that promote RNP biogenesis and engagement of the terminus (Cifuentes-Rojas and Shippen, 2012).

Arabidopsis thaliana is unusual in that it contains three TER isoforms. TER1 is a canonical telomere repeat template necessary for telomere length maintenance (Cifuentes-Rojas et al, 2011). In contrast, TER2 is a novel negative regulator of telomerase activity (Chapter II). The two RNAs assemble into distinct RNP complexes *in vivo* with different protein composition. Negative regulation by TER2 appears to proceed in part by competitive inhibition *in vitro*. Unlike TER1, TER2 does not serve as a template for telomerase action at chromosome ends *in vivo*. *In vitro* binding studies show that TERT binds ~10-fold tighter to TER2 than to TER1, arguing that TER2 can outcompete TER1 for the catalytic subunit, and then sequester this component in a non-

productive complex. TER2 is expressed at ~10-fold lower levels than TER1 under standard growth conditions. However, TER2 levels rise in response to DNA damage, so that TER2 becomes the predominant TER2 isoform (Chapter II). Concomitantly with increase of TER2, telomerase activity declines. Most importantly, telomerase inhibition in response to DNA damage is dependent on TER2, as a null mutant in TER2 abolishes DNA damage induced telomerase repression. Thus, TER2 is a novel non-coding RNA that modulates telomerase activity in response to DNA damage.

TER2 contains two regions of high similarity to TER1 (conserved regions 1 and 2; CR1 and CR2) that are separated by a 529nt intervening sequence (IS) which bears little similarity to TER1. TER2s is processed *in vitro* and *in vivo* to generate TER2s, a truncated RNA in which the IS is removed and the 3' 36nt are removed. TERT has much lower affinity for TER2s (~15-fold), suggesting the IS and/or 3' tail in TER2 is responsible for the tighter association with TERT (Chapter II). *In vivo*, preliminary evidence suggests that splicing and 3' end cleavage may be regulated by the DNA damage response (Chapter II). Thus, processing of TER2 appears to play an important role in telomerase regulation.

Canonical splice sites are absent from the IS. Instead, IS removal appears to proceed by a novel autocatalytic mechanism. *In vitro*, the IS is removed in a nucleotide-independent manner reminiscent of Group II self-splicing introns (Hernandez A, Cifuentes-Rojas C, and Shippen D, unpublished data). However, Group II introns are much larger that the TER2 IS (430-670bp in plants) and often contain an ORF encoding for a reverse transcriptase (Zimmerly et al, 2001), which is absent of TER2 IS. Thus far,

Group II introns have only been associated with organelles in eukaryotes (Dai et al, 2003). Finally, plant organellar Group II introns have largely lost their self-splicing capabilities, and their processing is dependent on nuclear-encoded maturases (Lambowitz and Zimmerly, 2004, Lambowitz and Zimmerly, 2011; Malek et al, 1997; Bonen L, 2008; Keren et al, 2009). The only known active ribozymes in the nuclear genome of plants are the hammerhead class, which promote RNA cleavage and not splicing (Przybilski et al, 2005). Thus, the TER2 IS has unique characteristics that do not fit with the class of canonical self-splicing introns.

The TER2 IS in *Arabidopsis thaliana* col-0 ecotype has enabled TER2 to become a novel competitive inhibitor of telomerase (Chapter II). Here we take a systematics and molecular approach to explore the nature and ancestry of the TER2 IS. We make use of the wealth of information provided by the 1001 Arabidopsis genomes project to examine the conservation of the TER2 IS in 513 distinct ecotypes of *A. thaliana*. We show remarkable divergence within the TER2 IS, and provide evidence that the TER2 IS has similarities to a broader group of transposable elements found within multiple members of the Brassicaceae family.

Materials and Methods

Sequence acquisition and analysis

Sequences corresponding to TER2 (Genbank accession number: HQ401285.1) were obtained using the genome browser at http://signal.salk.edu/atg1001. The search query AT5G24660 was used to pinpoint the region of interest, and all available tracks

(ecotypes) were selected. Two sequences were removed from our analysis. The ecotype Hov 3-2 was removed because it was the only ecotype with two deletions in the 5' end, corresponding to 20 nucleotides from the 5' start of TER2, and a 100nt deletion starting at nucleotide #101. The template region was not disturbed in this ecotype, possibly indicating a functional TER2 is generated. The Tottarp-2 ecotype was removed because the sequence corresponding to our search region did not contain a TER2. Importantly, it also did not contain a template region.

Sequences were trimmed in MEGA5, and then analyzed using Geneious (Drummond et al, 2010). Sequence conservation and alignments were performed using Geneious. IS-like sequences were obtained by BLAST searches of the *A. thaliana* (www.arabidopsis.org), *A. lyrata*, *Capsella rubella*, *Brassica rapa*, and *Thellungiella halophila* genomes accessed via www.phytozome.net (Hu et al, 2011; JGI; Cheng et al, 2011; *Thellungiella halophila* Genome Project 2011).

RNA extraction and QRT-PCR

RNA was extracted from 10-day old *Arabidopsis thaliana* ecotype Col and Ler-0 seedlings using the EZNA RNA extraction kit (OmegaBiotek) according to manufacturer's guidelines. Reverse transcription was performed using the Superscript cDNA master mix (Quanta), according to the manufacturer's instructions. To determine TER2s levels in both of these backgrounds, Q-PCR was performed on a Bio-Rad CFX-1000 using the following primers: TER2s F1: 5'-TACGGCAACAGAACCAGAGA-3'; TER2s R1: 5'-CTCCGACGAGACGACCATAC-3'. GAPDH primers: . Data was analyzed using Bio-Rad's CFX manager software. ΔΔCT values were obtained by comparing against GAPDH levels. Primer efficiency was determined on a per-run basis using Linreg (Ramakers, et al, 2003).

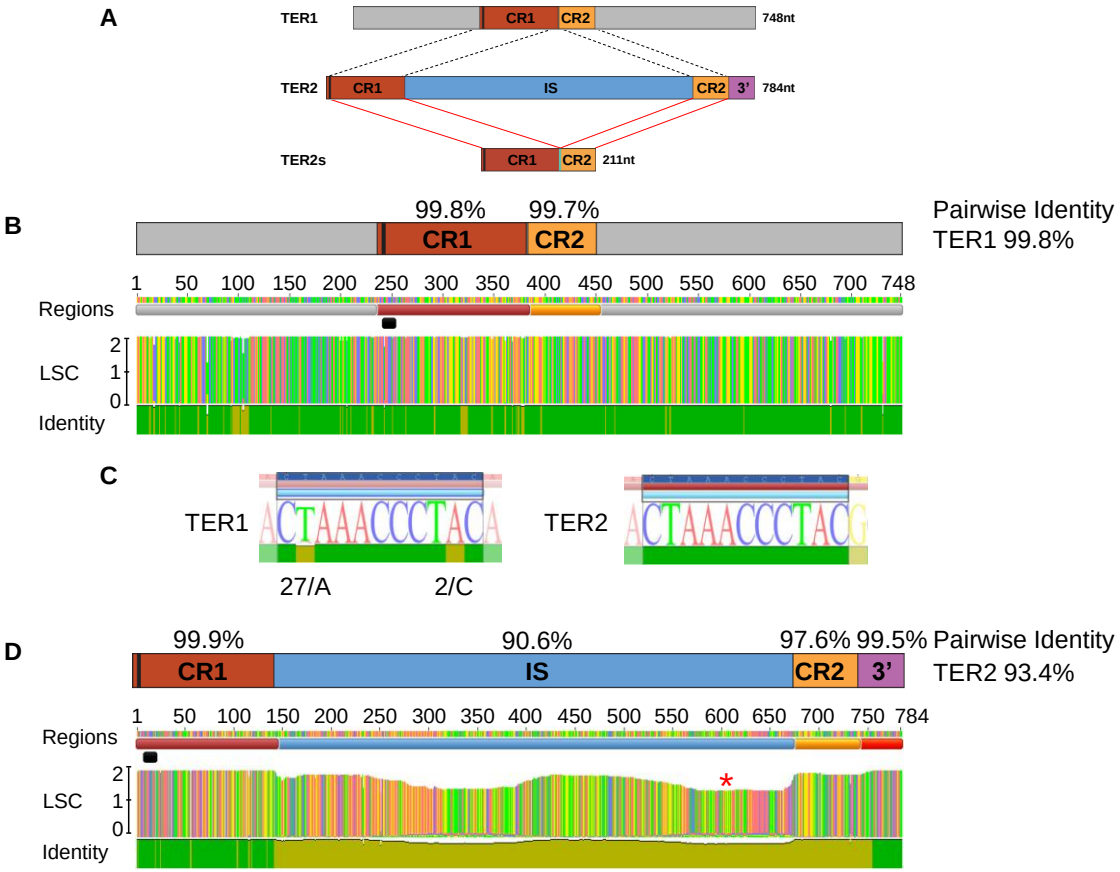


Fig 3-1. Comparison of different regions of TER2 and TER1 in 511 *A. thaliana* ecotypes. (**A**) Schematic diagram showing conserved regions between TER1, TER2, and TER2s. TER1 and TER2 share a core region of ~219nt comprised of conserved regions 1 and 2 (CR1 and CR2). The telomere template is denoted by a vertical black bar in CR1. TER2s is formed by two processing steps: 1) a splicing reaction that removes the IS and 2) a cleavage reaction that eliminates the 3' terminus (3' R). (**B**) Analysis of TER1 conservation between 513 different *A. thaliana* ecotypes. Identity shown in green denotes regions 100% conserved throughout all ecotypes whereas mustard yellow indicates variation. Local sequence confidence (LSC) is in a log base 2. LSC of 2 indicates a nucleotide was observed 100% of the time at that location. A reduction in this factor indicates a certain percentage of deletions at this site (example shown by the red asterisk in (**D**)). A green bar for identity corresponds to a LSC of 2. Pairwise identity for each region is denoted in % above each RNA region or for the entire RNA to the right. The telomere template region corresponds to the horizontal black bar. (**C**) Enlargement of the template region from both TER1 and TER2. The number below TER1 (left) denotes mutation frequency (#/511) and the nucleotide that was found at that site.
(**D**) Analysis of TER2 across all 511 ecotypes.

Results

TER2 conservation within Arabidopsis thaliana

The TER1 and TER2 locus from *Arabidopsis thaliana* were obtained from 511 of 513 available ecotypes (http://signal.salk.edu/atg1001), and then analyzed for variation against the Columbia ecotype. The template and 3' end of TER1 are embedded within the 5' UTR of a recently characterized Rad52 ortholog (Samach et al, 2011). TER1 is highly conserved among the different ecotypes, even the 5' region of TER1 that lies upstream of Rad52. (Figure 3-1A, and Figure 3-1B) TER2 is comprised of four distinct regions: two short regions that show high conservation with TER1 (CR1 and CR2; 144 and 67 nts, respectively), separated by a 529nt unique intervening sequence (IS), and terminated by a 36nt 3' region (3'R) (Figure 3-1A). CR1 retains very high identity between the ecotypes (99.7%), whereas CR2 shows slightly less nucleotide conservation (97.6%). The template region within TER2 is found in CR1, and as expected, shows 100% conservation between the different ecotypes (Figure 3-1C, right, and 3-1D).

Intriguingly, the template region of TER1 showed variation in two positions (Figure 3-1C, left). The A-C transition occurred only twice and potentially could reflect sequencing errors. The T-A variation was observed in 27/511 ecotypes. Nucleotide changes at these positions affect fidelity of telomerase translocation and nucleotide addition. How these alterations within the TER templating domain affect telomerase activity in ecotypes other than Col-0 is unclear.

The only other region of TER2 showing extremely high pairwise identity is the 3'R (99.5%). This finding is unexpected since this region of the RNA is removed from

the final spliced TER2s product and lies within an intergenic region. For comparison, TER1 is 91.6% identical (99.8% pairwise identity) across the entire RNA (Figure 3-1B). This high level of conservation implies that CR1 and 3'R are essential for function.

The intervening sequence within TER2 is not retained in all A. thaliana ecotypes

Out of 511 ecotypes examined, 63/511 (12.25%) showed either complete or partial disruption of TER2 IS (Figure 3-2A). The complete absence of the IS was rare, accounting for only 3/63 examples (Figure 3-2A; ecotypes Ler-0, Baa, and No-0). To ensure that this locus is still active without the IS, we extracted RNA from both Ler-0 and No-0 and performed QRT-PCR. A RNA corresponding to TER2s was present in equal levels compared to TER2s from the Col-0 background (Figure 3-2B). This suggests that transcriptional regulation and retention of TER2s is not altered in ecotypes without the IS.

Within TER2 IS, there are two islands of ≥50% identity corresponding to 63nt IS conserved region 1 (IS-CR1), and a 123nt conserved region 2 (IS-CR2). Sequences flanking IS-CR2 (HV-1 and HV-2; Figure 3-2C) show hyper-variability within the 63 ISD ecotypes. Even in ecotypes with high conservation in the IS, 11/408 showed variability in one particular tri-nucleotide rich region within HV-1. Despite high rates of IS disruption, these all of the ecotypes still retain high similarity within CR-1 and the 3'R, underscoring the potential for biological significance in these regions.

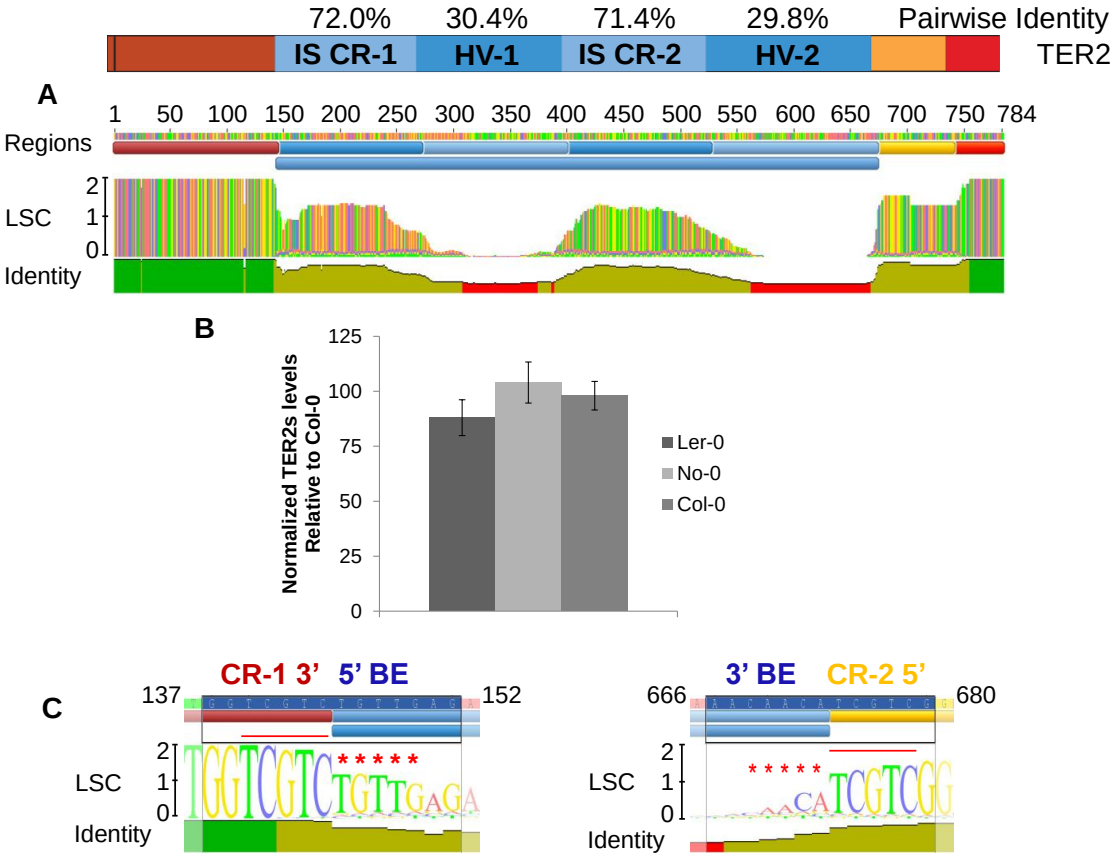


Fig 3-2. Comparison of the ISD ecotypes. (**A**) Analysis of the 63 ISD ecotypes, with regions of interest denoted. Pairwise identity recorded above the different IS regions mentioned in the text. Identities represented in red illustrate disproportional deletions (see red asterisks). (**B**) TER2s levels in Ler-0 and No-0. Fold change compared to Col-0, normalized to GAPDH. (**C**) 5' and 3' regions of the IS. The 5' (left) and 3' (right) regions of the IS have been enlarged to illustrate the inverted repeat found at the 5' and 3' boundary elements (BE) of the IS (red stars). A target site duplication is also present (TCGTC) at the 3' end and 5' beginning of the conserved regions 1 and 2, respectively (red bar).

We found that the 5’ and 3’ boundary elements (BE) of the IS (IS-BE) consist of inverted repeats (TGTTG/ACAAC). Strikingly, ecotypes with an imperfect match to the inverted repeat sequence within the 5’ IS-BE contained an IS disruption. Mutation or loss of the 5’ IS-BE most likely did not lead to IS disruption, as 41/63 ecotypes with an IS disruption retained the 5’ IS-BE (Figure 3-2C). Disruption of the 3’ IS-BE was more highly correlated with IS disruption, as only 19/63 of the ecotypes showed full 3’ IS-BE conservation.

The 3’ end of CR-1 and 5’ end of CR-2 also share a repeat sequence (TCGTC) at both sites (Figure 3-2C). In contrast to the 5’ IS-BE, only 2/63 (5/511) show any nucleotide variation at this at this site. Aside from the 5’ and 3’ ends of TER2, this region contains the only absolutely conserved nucleotide. The 5’ end of CR-2 is slightly less conserved, with 11/63 (13/511) ISD ecotypes showing any polymorphisms at this site (Figure 3-2A).

Multiple copies of the IS are present in A. thaliana

A BLAST using the TER2 IS as a query returned one full-length result on the left arm of chromosome 3 bearing 94.6% identity termed TER2 intervening sequence like (ISL-1, adjacent to At3G30120), and another hit on the right arm of chromosome 3 showing only 63.4% identity to TER2 IS (ISL-2, adjacent to At3G50210) (Figure 3-3A and 3-3B). Comparing these two ISL regions between the different ecotypes revealed two small conserved regions, roughly corresponding to IS-CR1 and IS-CR2 of the TER2 IS (Figure 3-3C; Figure 3-2A). ISL-1 and ISL-2 are found within intergenic regions and display a high number of single-nucleotide polymorphisms.

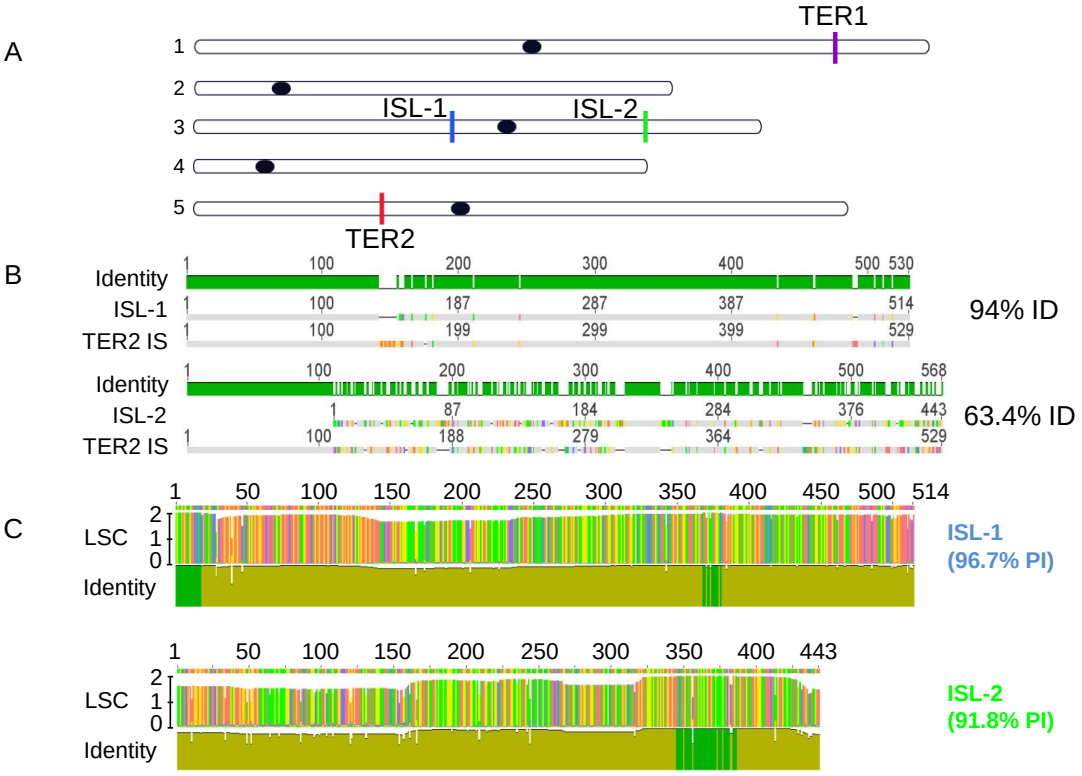


Fig 3-3. Three TER2 IS-like in *A. thaliana*. (**A**) Map of the five chromosomes in *A. thaliana* Col-0 with TER2 IS, ISL-1, ISL-2, and TER1 loci shown (adapted from TAIR). (**B**) Geneious global alignment of TER2 IS and ISL-1 (top) and ISL-2 (bottom). Percent nucleotide identity for the only regions that aligned are shown to the right (ie, 1-443 of ISL-2). (**C**) IS region comparisons between the 510 different *A. thaliana* ecotypes. Pairwise identity is shown to the right. Labels are color coded to correspond to the location in (**A**).

The IS shares similarities to Class II DNA transposons

As discussed above, analysis of the TER2 IS and the ISL elements within the different ecotypes reveals several highly conserved elements within the IS and in the sequences flanking them. One additional conserved sequence is a five nucleotide tandem inverted repeat (TIR) found at the 5' and 3' boundaries of both TER2-IS and ISL-1 (Figure 3-4A and 3-4B). Target site duplications (TSDs) also flanked all three elements (Figure 3-4A-C). TSDs ranged in length from 5nt for TER2-IS and ISL-1 to 18nt for ISL-2. The TSD nucleotide sequence varied, suggesting that these insertions were not the result of gene duplication, but rather represented unique insertion events.

The conserved boundary elements associated with TER2 IS and the ISL regions closely resemble Class II DNA transposable elements (TEs) (Reviewed in Feschotte et al, 2002). Class II TEs comprise ~8% of the *A. thaliana* genome and range in size from ~ 100bp to several kb (Hu et al, 2011; Feschotte and Pritham, 2007; Jiang et al, 2003). Of particular note are the miniature inverted-repeat transposable elements (MITEs). MITEs are typically short (100-600bp), do not contain ORFs encoding a DNA transposase, retain structural homogeneity within families, and show a preference for insertion near genes (Bureau et al, 1992; Feschotte et al, 2002; Zhang et al, 2000).

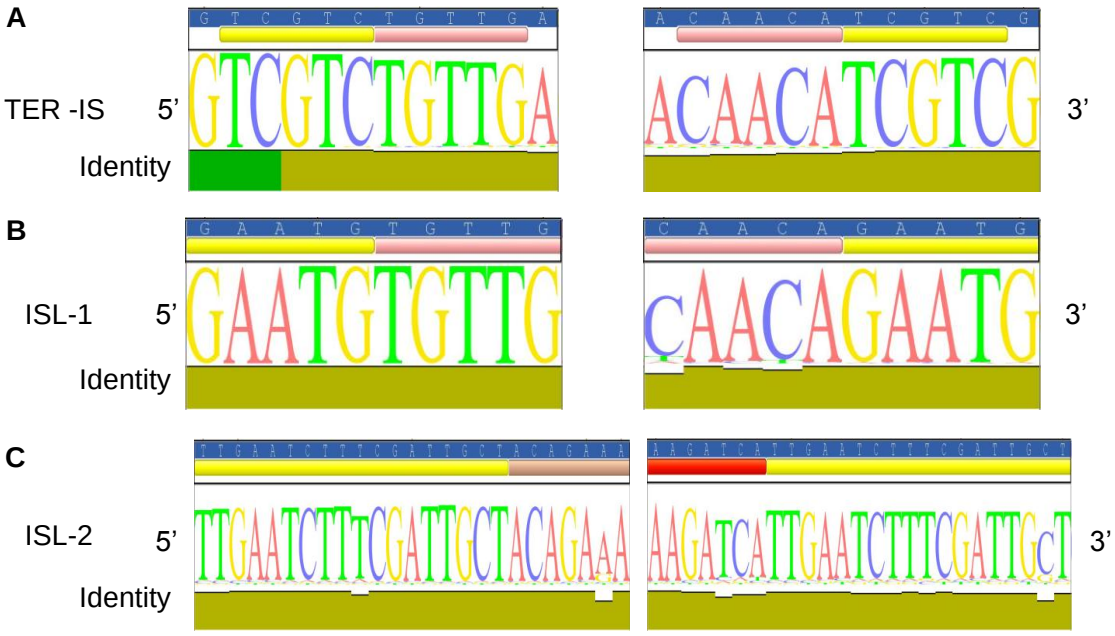


Fig 3-4. The 5’ and 3’ boundary elements of the IS elements in *A. thaliana*. Sequence immediately adjacent to the 5’ (left) and 3’ (right) boundaries of the TER2-IS (**A**), ISL-1 (**B**), and ISL-2 (**C**) are highlighted. Target site duplications (TSD) are shown in yellow. Tandem inverted repeat (TIR) are in red. Region of 100% sequence identity are indicated by the bright green bar.

The IS is present in multiple species of the Brassicaceae lineage

The *A. thaliana* genome is more compact (~125Mb) than its close relative *Arabidopsis lyrata* (207Mb) (Hu et al, 2011). In many organisms, including *A. lyrata*, TE abundance correlates with this increased genome size (Hu et al, 2011). Approximately 30% of the *A. lyrata* genome is comprised of TE, compared to ~24% in *A. thaliana* (Hu et al, 2011). Moreover, the genomic sequence devoted to TE has doubled in *A. lyrata* compared to *A. thaliana* (62:30Mb) (Hu et al, 2011). If the TER2 IS is associated with a TE, we predict it would be present more than three times in *A. lyrata*. Moreover, the ISL elements should be located in non-syntenic regions of the genome, corresponding to an expansion of the TE family. To test this prediction, we performed BLAST against sequenced genome databases from several members of the Brassicaceae phylogeny (Figure 3-5A). We first examined the *A. lyrata* genome, using the IS consensus sequence from *A. thaliana*. Hits were defined as returns longer than 100bp, bearing 70% identity to a corresponding region with the query IS. BLAST returned 32 unique hits dispersed throughout the *A. lyrata* genome. Two of these hits were "full length," retaining both the 5' and 3' 5nt TIR from *A. thaliana* (Figure 3-5B). Two others showed high similarity throughout the IS, but were lacked 90-130nt from the 3' end. A significant fraction (24/32) showed high similarity within the 5' 200nt of the IS (Figure 3-5B), including 9 examples that retained the 5nt TIR from TER2-IS (TGTTG). Notably, the *A. lyrata* IS elements localized on 10 different scaffolds and tended to cluster in groups 2-3kb apart from one another. This observation supports the

idea of IS family expansion. The insertion site for the youngest (most conserved) IS from scaffold 8 did not reveal any conserved insertion site requirements.

We also performed BLAST search against the ~250Mb *Capsella rubella* genome (Johnston et al, 2005). The split between the branches leading to Arabidopsis and Capsella likely occurred between 10-20mya (Figure 3-5A) (Beilstein et al, 2010). Therefore, *C. rubella* represents the next closest evolutionary relative to *A. thaliana*. Roughly 60% of the C. rubella genome has been constructed and made available. Using the consensus sequence of the youngest (full length) IS in *A. lyrata*, we identified three IS-like sequences within *C. rubella* that matched our previous search criteria (Figure 6-5B). These three IS-like regions, like their counterparts within *A. lyrata*, showed high similarity within the 5' 200nt and were positioned on two different scaffolds. BLAST analysis of *Brassica rapa* failed to reveal a sequence hit. However, one ISL element was identified in the equally divergent *Thellungiella halophila* genome (both diverged ~50mya) (Figure 3-5B and 3-5C). Analysis of *Carica papaya*, which diverged from *A. thaliana* ~100mya, failed to return any IS-like hits (Figure 3-5A). Altogether, these findings demonstrate that the TER2-IS shares similarity with transposable elements, specifically Class II TE, and further that this element is present throughout the Brassicaceae lineage.

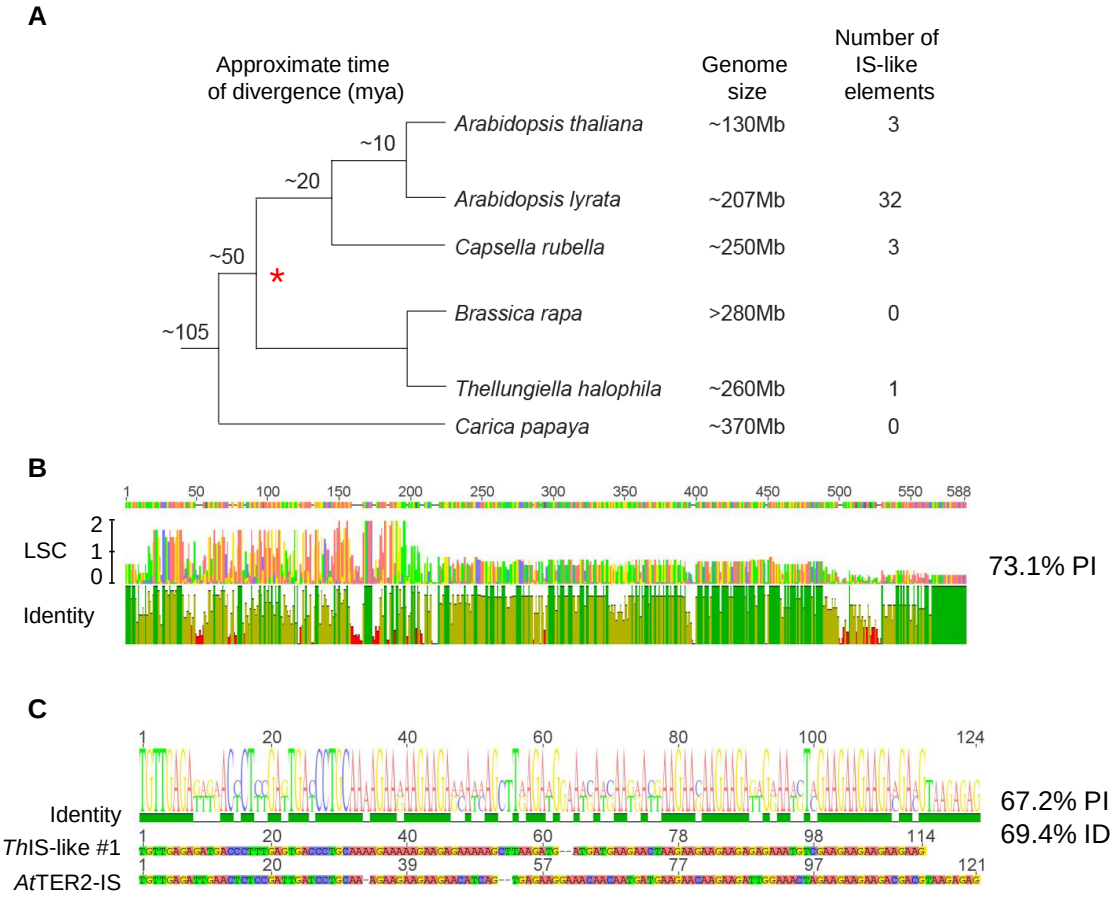


Fig 3-5. The TER2-IS is present in other members of the Bracicaceae family. (**A**) Phylogenetic tree of Brassicaceae family members (including the Brassicales member *Carica papaya*) that were analyzed for the presence of the IS element. Approximate time of divergence was adapted from (Beilstein et al, 2010). Red asterisk indicates the predicted last common ancestor to contain an IS element. Source of genome size estimates are listed in the text. (**B**) Alignment of the 32 potential IS-like sequences from *A. lyrata*. In this instance, local sequence conservation (LSC) (e.g. height of the colored line indicates regions of higher coverage (5' end). Only 5 elements spanned the 3' end (low coverage), whereas 24 covered the 5' end (high coverage). Thus, the 100% identity at the 3' end could be artificial, due to low numbers. (**C**) Geneious global alignment of the AtTER2-IS and the *Thellungiella halophila* potential IS-like element. Pairwise and nucleotide identity shown to the right.

Discussion

As we enter the era of population genomics, it is becoming easier to survey gene variation within a species, and from this data determine regions necessary for function and regulation. Because telomerase RNAs show a high degree of sequence variation between species, we took advantage of the 1001 Arabidopsis genomes project to define regions within TER1 and TER2 that were most conserved, and presumably most important for function.

TER1, the telomerase RNA subunit necessary for telomere maintenance in *A. thaliana*, was highly conserved in all of the ecotypes. Some of this conservation can be attributed to its position with the 5' UTR of a gene encoding RAD52 (Samach et al, 2011). TER1 conservation is not restricted to the protein coding region. The 5' half of TER1, which is does not overlap with the RAD52 ORF, shows high levels of conservation. This observation is intriguing, since the 5' UTR of putative Rad52 genes in the *A. lyrata* genome retain little similarity with TER1 and do not contain a template. Thus, TER1 is not present at the corresponding locus in *A. lyrata*.

While high conservation was also associated with the TER2 locus in the analyzed ecotypes, it was limited to the two regions shared with TER1 (CR-1 and CR-2), and the 3' terminus, implying these regions are required for the function of TER2. In contrast, the TER2 IS, which conveys TER2's inhibitory nature in Col-0, is remarkably divergent among different ecotypes. In three ecotypes the IS is completely absent from TER2. A phylogeny of the *A. thaliana* ecotypes suggest these three ecotypes (Baa, Ler-0, and No-0) are too divergent from one another to represent an ancestral, pre-IS state. Instead, this

implies that they each lost the IS independently of one another. (http://signal.salk.edu/atg1001/index.php). Thus, unlike the rest of TER2, the TER2 IS is in a state of flux.

Preliminary *in vitro* characterization of the self-splicing reaction of TER2 show that the 3' end is necessary for processing. However, nothing is known about the requirements within the IS itself. The 3' end of TER2 is completely conserved, yet the IS is highly variable in 63/511 ecotypes, raising the question of whether splicing occurs in ecotypes with large IS deletions. It is also unclear what effect IS variability has on inhibition of telomerase by TER2. Answering these questions requires detailed biochemical probing of the IS. Nevertheless, the current data suggest that inhibition by TER2 is not necessary for fitness, but that the function performed by TER2s is critical.

At least three copies of the TER2 IS are present in the *A. thaliana* genome. The TER2 IS shows the highest degree of variation of the three, possibly reflecting on the age of these elements, or a selective pressure to "inactivate" the IS within TER2. Among the three IS elements, we found regions conserved between them reminiscent of Class II DNA transposons, including tandem inverted repeats (TIRs) and target site duplications (TSDs). All three IS elements are found adjacent to predicted protein-encoding genes. In addition, these elements are found throughout the Brassicaceae lineage, expanding in number in *A. lyrata*. These attributes, plus the small size of the IS (443-529nt), suggest that the IS belongs to a novel class of miniature inverted transposable elements (MITEs).

TER2 IS and the ISL elements are highly conserved, but few in number in *A. thaliana*, consistent with a young but passive family of TEs. Our data suggest that the

TER2 IS family has been much more active in *A. lyrata*, as it includes at least 32 members, some with very high conservation throughout the entire length of the element. In addition, regions adjacent to full-length IS-like elements show several hundred bp of similarity, consistent with transduplication. Non-TE DNA is duplicated along with the TE (Le et al, 2000; Jiang et al, 2004; Hoen et al, 2006). Evidence for transduplication was observed multiple times in *A. lyrata*, with the longest non-TE tract of DNA corresponding to ~450bp. While further analysis is required to delineate the boundaries and age of the TER2 IS elements, our data indicate that these elements have novel characteristics. If the duplicated non-TE DNA we observed arose from a transduplication event, then the IS may have had some role in the duplication of the TER gene in *A. thaliana*.

TEs frequently mobilize during periods of organismal stress, shuffling neighboring genes in the process (Mirouze and Paszkowski, 2011; Slotkin and Martienssen, 2007). In some situations, transposon shuffling is favorable. The multiple TEs that are adjacent to *teosinte branched1* (tb1) gave rise to domesticated maize (Studer et al, 2011). Further analysis of TER2 IS may not only reveal insight into the dynamic nature of telomerase regulation, but also the resilience of all genomes to make use of potentially detrimental “junk DNA.”

CHAPTER IV

*AT*POT1B: A TELOMERASE ACCESSORY FACTOR AND CRITICAL CAPPING COMPONENT IN *ARABIDOPSIS THALIANA*

Summary

Telomeres consist of tandem arrays of DNA repeats at the end of linear chromosomes that resemble double-strand breaks. Thus, it is tantamount that telomeres be constantly protected and sequestered away from DNA repair pathways. One of the conserved protein complexes that bind and protect telomeres from illegitimate activities is the six-membered complex shelterin. In the shelterin complex, one protein, POT1, binds directly to ssDNA. This binding, in conjunction with the rest of the shelterin complex, represses the DNA damage response pathways and regulates extension of the chromosome-end by the telomerase RNP complex. Therefore, telomere length maintenance and chromosome-end protection is a carefully regulated process. In the model plant *Arabidopsis thaliana*, multiple duplication events have give rise to three POT1 proteins. Previous characterization has determined that POT1a and POT1b are both components of distinct telomerase RNP complexes. POT1a binds to one of the three TERs in Arabidopsis, TER1, and is required for telomere homeostasis and robust telomerase activity. POT1b binds to TER2, the TER component of an inhibitory telomerase RNP complex. Here I characterize a mutant background lacking POT1b, and show that this line, *pot1b-1*, displays very few telomere defects, but shows increased telomerase activity, consistent with a role in telomerase regulation. Strikingly, crosses of

pot1b-1 and a TER2 mutant background, *ter2-1*, results in profound genome instability and physiological defects. Thus, preliminary evidence suggests POT1b and TER2 are part of a novel telomere capping complex.

Introduction

Telomeres and their protein constituent provide a protective cap on the chromosome end that prevents it from being recognized as a double strand break (DSB) and illiciting an inappropriate DNA damage response. Telomeres are comprised of tandem, G-rich repeats terminated by a short single-strand (ss) 3' overhang, termed the G-overhang. Two major telomere protein complexes appear to define the chromosome terminus. The six-member shelterin complex was first identified in vertebrates, and is made up of two double-strand (ds) DNA binding proteins, TRF1 and TRF2, the TRF2 interacting protein RAP1, a ssDNA binding heterodimer POT1/TPP1, and finally, TIN2, which acts as a bridge between the ds and ss regions. In budding yeast, the G-overhang binding heterotrimer Cdc13/Stn1/Ten1, also known as CST, associates with the chromosome terminus. In addition to protecting telomeres, Shelterin and CST facilitate telomere length maintenance. Telomeres act as a buffer against the end-replication problem, and must periodically be lengthened to replace DNA lost due to the properties of semi-conservative replication.

Telomere length maintenance is a highly regulated process, centered on telomerase. Telomerase is ribonucleoprotein complex minimally comprised of a reverse transcriptase TERT and an RNA subunit called TER. TER provides a template for

synthesis of G-rich telomere repeats by TERT. In most organisms, TERT and TER are encoded by a single copy gene. However, in the flowering plant, *Arabidopsis thaliana*, two TER subunits are encoded, TER1 and TER2 (Cifuentes-Rojas et al, 2011). TER1 functions as a canonical telomerase template and is required for telomere length maintenance. TER2, by contrast, is a novel negative regulator of telomerase activity that appears to out-compete TER1 for TERT, and then sequester the catalytic subunit in a non-productive complex (Cifuentes-Rojas et al, 2011; Chapter II). Notably, TER2 is further processed *in vivo* into a third TER isoform, TER2s. TER2s is produced by removal of a 529nt intervening sequence between the two conserved regions, R1 and R2, and the cleavage of 36nt from the TER2 3' end. The final 219nt TER2s shows 85% identity to the corresponding region of TER1.

While all three TER isoforms reconstitute telomerase activity *in vitro* (Cifuentes-Rojas et al, 2011), they perform different functions *in vivo*. In flowers, where telomerase is highly active, TER1 and TER2s are present at almost equivalent levels. In contrast, TER2 is approximately 10-fold less abundant than either of the other TER isoforms in all tissues, including vegetative organs were telomerase is repressed. Over-expression of TER2 *in vivo* leads to a shift in the ratio of TER2:TER2s, with TER2 becoming the predominant isoform. Moreover, plants overexpressing TER2 exhibit reduced telomerase activity and progressive telomere shortening. Thus TER2 is a negative regulator of telomerase activity. TER2 levels are rapidly up-regulated in response to DNA damage, resulting in repression of telomerase activity (Chapter II). Moreover, DNA-damaged induced telomerase repression is dependent on TER2.

Therefore, TER2 is a novel DNA damage induced non-coding RNA that modulates telomerase activity. The function of TER2s is unknown.

Aside from the core components TERT and TER, several accessory factors associate with telomerase and play roles in RNP biogenesis, enzymology, and engagement with the chromosome termini (Collins K, 2006). In budding yeast, Est1p facilitates telomerase recruitment through interactions with CST and the yeast TER, Tlc1. The interaction between Est1 and Cdc13 is necessary for telomerase to act on telomeres in late S-phase. However, telomerase can be detected at telomeres throughout the cell cycle in an Est1-independent manner. This recruitment is dependent on the non-homologous end-joining factor Ku, which interacts with DNA, as well as Tlc1, although not simultaneously. In vertebrates, the mechanism of telomerase recruitment is unknown, although, the Shelterin components TPP1 and POT1 are implicated in this process (Tejera et al, 2010).

The POT1 (Protection of Telomeres) family of proteins in partnership with TPP1, form an ancient and functionally conserved telomere-specific complex (Baumann and Cech, 2001). Functional orthologs of this heterodimer have been identified in many metazoan species, including fission yeast, ciliates, and vertebrates (Baumann and Cech, 2001). The POT1/TPP1 heterodimer binds ss telomeric DNA, sequestering it from potential exonucleolytic activity and recognition by DNA damage sensing complexes found at telomeres. Loss of POT1 leads to telomere elongation and activation of a severe DNA damage response (Baumann and Price, 2010).

POT1 structural homologs have been identified in many species. Intriguingly, POT1 copy number and function varies. *POT1* has been duplicated in several species, including mice, Tetrahymena, *C. elegans*, and many plant species. Duplication events provide the potential for species-specific variation or sub-functionalization in function. For example, the murine POT1 genes have undergone a duplication and subsequent subfunctionalization (Palm et al, 2009). MmPOT1a represses ATR activation and prevents NHEJ at newly replicated telomeres, whereas MmPOT1b regulates C-strand resection events (Kibe et al, 2010). Tetrahymena POT1 has also duplicated, but only POT1a is essential for telomere length maintenance (Jacob et al, 2007). C. elegans telomeres are protected by at least two POT1-like proteins that protect either G or C-rich overhangs (Raices et al, 2008).

Many shelterin components, such as TPP1 and TIN2, are not found within plants. However, analysis of POT1 in the moss species *Physcomitrella patens* indicates that the telomere capping function of this protein family is conserved at the base of the plant lineage (Shakirov et al, 2010; Appendix A-2). A null mutation in *P. patens* POT1 results in severe developmental defects, shortened telomeres and increased G-overhangs. In both the monocot and dicot lineage, there is evidence for two independent POT1 duplication events, one in the Poaceae family, and another in Brassicaceae, which includes the model organism *Arabidopsis thaliana* (Shakirov et al, 2009; Beilstein et al, in preparation).

A. thaliana appears to be unusual among flowering plants in that it encodes three POT1 proteins, AtPOT1a, AtPOT1b and POT1c, but, like Physcomitrella and all other

plants, lacks an apparent TPP1 ortholog. Unlike the previously characterized POT1 proteins, AtPOT1a is a critical telomerase accessory factor (Shakirov et al, 2005; Surovtseva et al, 2007; Cifuentes-Rojas et al, 2011) that shows specificity for TER1 both *in vitro* and *in vivo*. A null mutation in POT1a results in a 15-fold reduction of telomerase activity and progressive telomere shortening (Surovtseva et al, 2007). POT1a is enriched at telomeres during S-phase, and *in vitro* binds to two components of the CST complex, CTC1 and STN1. POT1a displays evidence of positive evolutionary selection driving its interaction with the CST complex (Beilstein et al, in prep). Thus, POT1a may play a role in telomerase recruitment, similar to Est1p in budding yeast (Surovtseva et al, 2007; 2009).

Little is known about the role of AtPOT1b. Over-expression of the N-terminal portion of AtPOT1b, leads to rapid shortening of telomeres, chromosome fusions, and severe growth and developmental defects (Shakirov et al, 2005), consistent with an essential role for this protein in chromosome-end protection. Like AtPOT1a, AtPOT1b does not associate with telomeric DNA *in vitro* and instead binds to TER2 (Cifuentes-Rojas et al, 2011; Chapter II). AtPOT1b assembles with Ku into a TER2 RNP complex *in vivo* that is distinct from the TER1 RNP (Cifuentes-Rojas et al, 2011). Notably, POT1b does not display marks of positive selection like POT1a. Rather, POT1b more closely resembles single copy POT1 proteins from plants (Beilstein et al, in prep). Thus, it has been unclear how POT1b contributes to telomere biology. Is it necessary for chromosome end protection or is involved in telomerase regulation? Here we report that AtPOT1b regulates the processing of TER2. In the absence of AtPOT1b, TER2 is lost

and abberant RNA processing intermediates accumulate. Furthermore, we show that AtPOT1b works in concert with TER2 to protect chromosome ends. These data suggest an intricate and evolving relationship between telomerase and the telomere cap.

Materials and Methods

Plant materials and Flag-Myc-POT1b construction

The pot1b-1(Ler-0) line was isolated from a gene trap collection (CSHL). To genotype the wild type locus, P2GT1F: 5'-AAACCCCAACGATCAGAGAC-3' and P2GT3R: 5'-AGACGAAGAGGTTGTTTCATTGCA-3' were used. The mutant locus was determined using P2GT3R and DS3-1: 5'-ACCCGACCGGATCGTATCGGT-3'. The ter2-1 lines have been described previously (Chapter II). *pot1apot1b* crosses were generated by crossing lines harboring the pot1a-1 (Surovtseva et al, 2007) and pot1b-1 alleles. Plants were grown under standard conditions (16h-light/8h-dark). To obtain a Flag-Myc-POT1b transgenic line, full length POT1b cDNA was Gateway cloned into vector pB (kind gift of Z. Xiuren, TAMU) using LR clonase II plus (Invitrogen) and transformed into *Agrobacterium tumefaciens* strain GV3101. Transformation of *pot1b-1* and wild type plants was performed using the floral dip method. T1 transformants were selected on 0.5 MS media supplemented with 400mg/L glufosinate ammonium (BASTA, Crescent Chemical), genotyped, and analyzed by western blotting for transgene expression.

Q-TRAP, TRF analysis, In-gel G-overhang assay

Q-TRAP and TRF analysis were performed as described previously (Kannan et al, 2008; Surovtseva et al, 2009). G-overhangs were monitored by the in-gel hybridization technique described previously (Surovtseva et al, 2009) with the following modification. After drying the gel, hybridization was performed in a hybridization oven in a glass tube instead of in a temperature controlled water bath. Phosphoimaging was performed using a Bio-Rad Pharos FX phosphoimager and analyzed using QuantityOne software.

RNA extraction and cDNA synthesis for QRT-PCR

RNA extraction for QRT-PCR was performed using EZNA RNA extraction kit (Omega Biotek) according to the manufacturer's instructions. For reverse transcription, one microgram of RNA was added to the Quanta cDNA synthesis master mix, per their protocol. Quantitative PCR was performed with Bio-Rad's Sso Advanced SYBR on a Bio-Rad CFX-1000 instrument. To distinguish between different isoforms of TER2, the following primers were used: (QRT-TER2S F1 and R1) 5' tacggcaacagaaccagaga 3' and 5' ctccgacgagacgaccatac 3', (QRT-TER2 FL F and R) 5' tttgtacggcaacagaacca 3' and 5' gcaggatcaatcggagagtt 3'. See figure 4-3A for nomenclature. RNA levels were analyzed using the LinReg software (Ruijter et al, 2009) and normalized to GAPDH and TIP41L.

Results

Plants deficient for POT1b do not display a chromosome end deprotection phenotype

Because previous analysis of POT1b function involved over expression, which can lead to confounding dominant-negative effects, we sought to identify a null mutation

in *POT1b*. In the CSHL gene trap collection, there is an *A. thaliana* line with a T-DNA insertion within the second exon of POT1b, which we termed *pot1b-1*. RT-PCR failed to detect AtPOT1b transcripts in this line, indicating that it is a null mutant (Figure 4-1A, 4-1B). In contrast to the growth defects associated with POT1b over expression lines, pot1b-1 mutants were indistinguishable from wild type, except for being slightly less robust in stature (Data not shown). Terminal restriction fragment (TRF) analysis showed no difference in bulk telomere length for the pot1b-1 and wild type (Figure 4-1C). Telomeres ranged from 2-5kb in the wild type and pot1b-1 setting. In contrast, G_1 *pot1a-1* telomeres displayed a ~1kb loss and showed a banding pattern similar to *tert* mutants. We also failed to detect telomeric fusions using telomere fusion PCR or conventional cytological analysis of anaphase spreads. This result was not unexpected, given the wild type length of telomeres in this background. Loss of POT1 genes in other organisms typically results in deregulation of G-overhangs, leading to increased G-overhang signal. Therefore, we examined the status of G-overhangs in pot1b-1 by the native in-gel hybridization technique. We found that G-overhang signal was reduced by approximately 50% in the *pot1b-1* setting compared to wild type, whereas G-overhang signal was slightly increased in pot1a mutants (Figure 4-1D). We conclude that removal of POT1b does not grossly destabilize the chromosome end, and further that AtPOT1b has a function distinct from AtPOT1a.

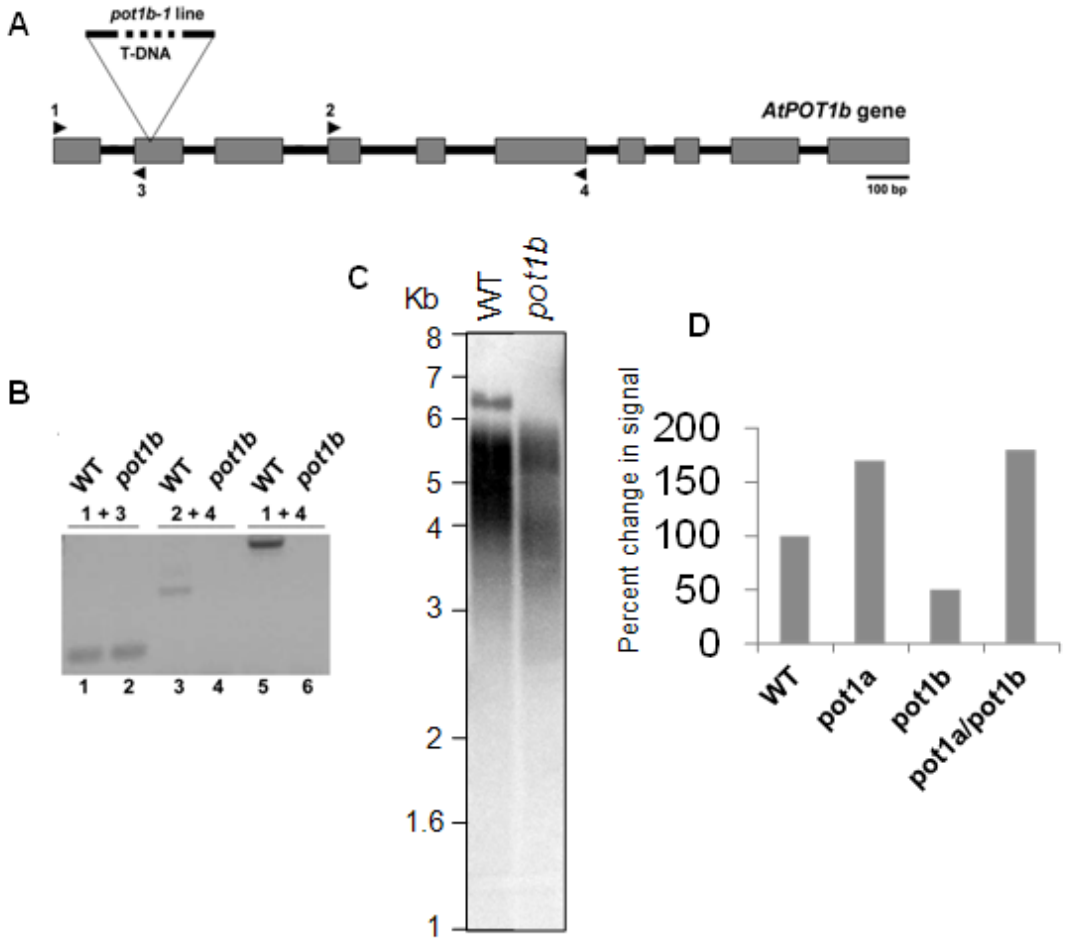


Fig 4-1. *Atpot1b* lines display a minimal telomere de-protection phenotype. (**A**) Diagram showing the gene structure of AtPOT1b. Exons are displayed as boxes, introns as lines. The site of T-DNA insertion is shown within the second exon. Numbered arrows denote primer positions used in (**B**) to ensure loss of transcript production in the T-DNA lines. (**B**) RT-PCR to amplify the POT1b locus in *pot1b* and wild types lines. (**C**) Terminal restriction fragment length (TRF) analysis examining bulk telomere length in wild type and *pot1b* lines. Southerns were probed with a radio-labeled $(TTTAGGG)_4$ probe. Molecular weight markers are listed to the left in Kb. (**D**) Quantification of the change in G-overhang signals in wild type, *pot1a*, *pot1b*, and *pot1a/pot1b* double mutants.

AtPOT1b negatively regulates telomerase activity

Given that AtPOT1b interacts with TER2, we next asked whether telomerase activity levels were altered in *pot1b-1* lines. Quantitative TRAP (Q-TRAP) was performed with extracts from flowers. Telomerase activity was increased by 3-4 fold, similar to what we observed with *ter2-1* flowers (Figure 4-2A; Cifuentes-Rojas et al, in prep). Notably, extracts from *pot1b-1* seedlings also showed an increase in telomerase activity, an effect that was not associated with *ter2-1* mutants. To ensure that this effect was not due to variations between the Columbia (Col-0, *ter2-1*) and Landsberg (Ler-0, *pot1b-1*) ecotypes, activity was tested for in Ler-0 POT1b +/+. No difference in telomerase activity was observed between Col-0 and Ler-0 ecotypes (Figure 4-2A).

To ensure that the increase in telomerase activity in *pot1b-1* mutants is indeed caused by a mutation in AtPOT1b, we performed genetic complementation with wild type AtPOT1b. *pot1b-1* lines were transformed with an N-terminal 6xFlag-Myc POT1b construct under the control of the strong 35S cauliflower mosaic virus promoter (Figure 4-2B). As expected, telomerase activity levels returned to wild type in the T1 transformants, but not in untransformed siblings (Figure 4-2C). Therefore, we conclude that AtPOT1b contributes to the negative regulation of telomerase in *A. thaliana*.

AtPOT1a is a TER1 telomerase RNP component that is essential for proper telomere length maintenance *in vivo (*Surovtseva et al, 2007; Cifuentes-Rojas et al, 2011). In addition to progressive telomere shortening, pot1a mutants display a dramatic decrease in telomerase activity (Surovtseva et al, 2007; Figure 4-2D). A genetic approach was taken to determine if POT1a and POT1b work in similar pathways to regulate telomerase activity. *pot1a-1/pot1b-1* double mutants were generated and examined for telomerase activity. Strikingly, although telomeres in this background are shorter and thus resemble those of pot1a-1 single mutants (data not shown), telomerase activity was increased by 2-3 fold compared to wild type, similar to pot1b-1 single mutants (Figure 4-2D). Thus, POT1b affects telomerase activity in a different manner from POT1a. In addition, the increased telomerase activity in the pot1apot1b double mutants is not sufficient to rescue the telomere maintenance defect of pot1a. This supports the model that POT1a is needed to recruit telomerase to the chromosome end.

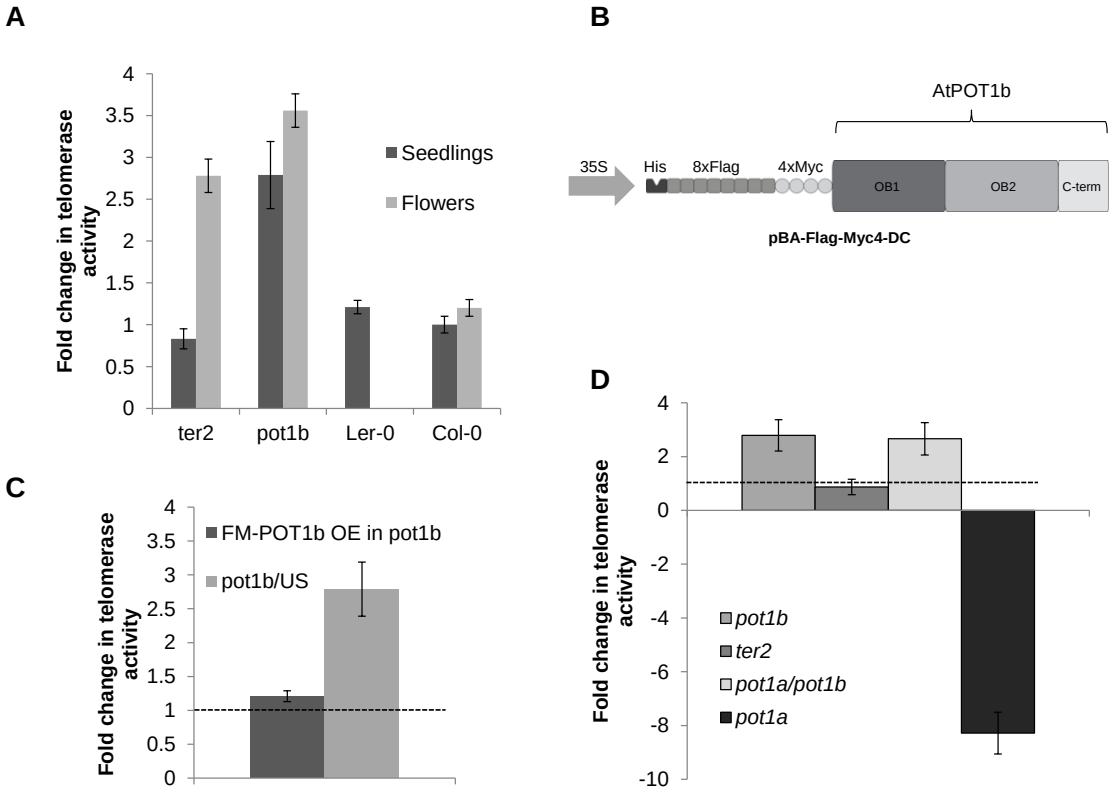


Fig 4-2. *pot1b* mutants show reduced telomerase activity similar to *ter2*. (**A**) QTRAP on extracts from *pot1b-1* and *ter2-1* seedlings and flowers. 14-day (post-germination) seedlings and flowers were tested for telomerase activity. n≥5 for each background. Error bars represent S.D. (**B**) Construct design for complementing *pot1b-1* lines. (**C**) QTRAP on positive transformants expressing the Flag-Myc-POT1b construct (FM-POT1b). Unselected *pot1b-1* transformants were used as a control (pot1b/US). n=3. (**D**) Telomerase activity in pot1a/pot1b seedlings. QTRAP was performed on 14-day seedlings. In (**C**) and (**D**), all fold changes are relative to wild type (dashed line).

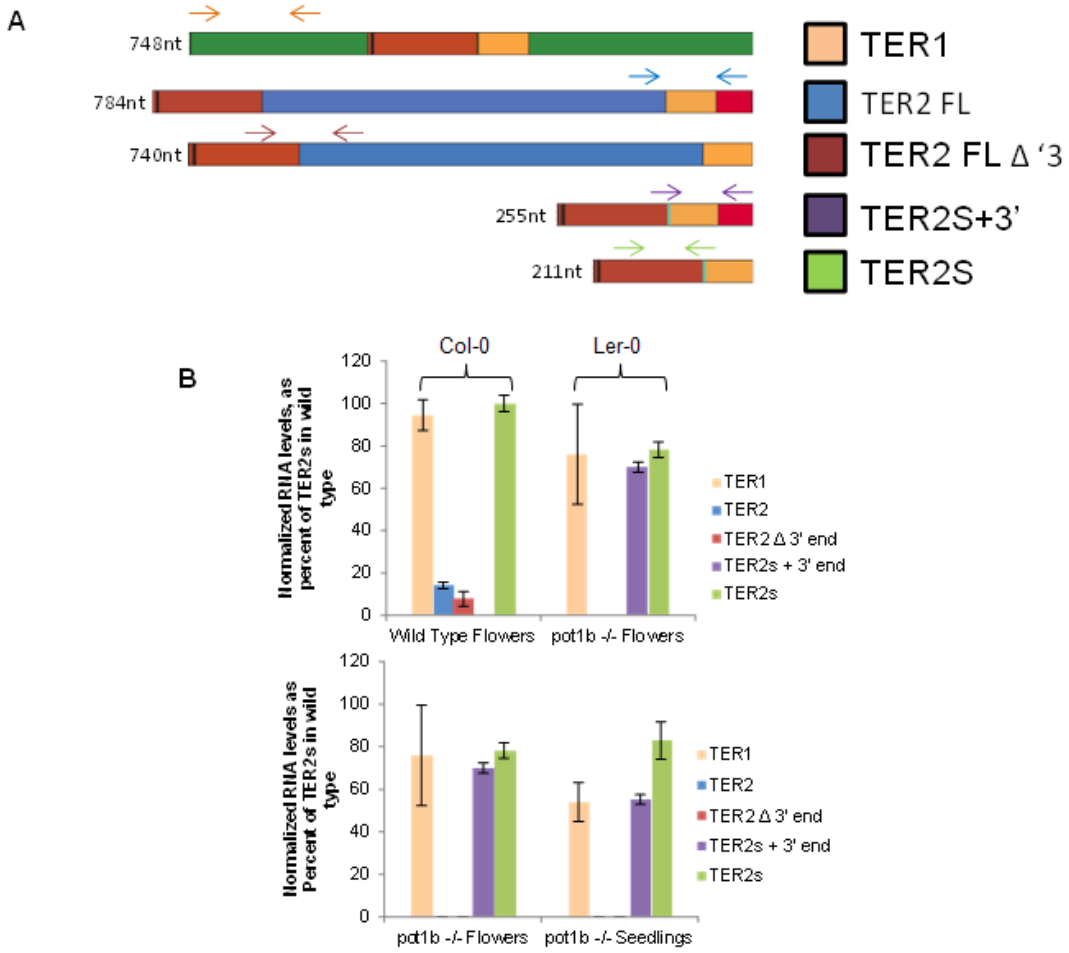


Fig 4-3. POT1b is necessary for the proper 3' end processing of TER2. (**A**) Schematic representation of the TERs in Arabidopsis. Dark red and orange boxes within TER1, TER2, and TER2s represent regions of ≥90% identity conservation between the RNAs. Arrows above each RNA indicate the region targeted by QRT-PCR primers. Color coded boxes to the right correspond to QRT-PCR products in (**B**). (**B**) TER levels in wild type and *pot1b-1* mutants. QRT-PCR was performed on RNA extracted from flowers (top) and seedlings (bottom).

AtPOT1b is critical for proper stability and splicing of TER2

One explanation for the overlap in phenotypes in *ter2* and *pot1b* mutants is that AtPOT1b stabilizes TER2. To test this hypothesis, QRT-PCR was performed on RNA from wild-type and *pot1b-1* flowers. Primers were designed to specifically amplify TER1, TER2, and TER2s, as well as putative TER2 processing intermediates (Figure 4-3A). We recently discovered that the intervening sequence (IS) within TER2 is absent in the Ler-0 ecotype, which is the background of *pot1b-1* (Chapter III). While the peculiarities of the IS are discussed in Chapter III, because the Ler-0 TER2 is similar in sequence and expression to TER2s from Col-0 it will be referred to as TERs. Due to this nuance of Arabidopsis variation, we focused on the role of AtPOT1b in TER2s processing and maintenance.

In wild-type Col-0 and Ler-0 seedlings, TER1 and TER2s were present at equivalent levels (Figure 4-3B), as previously observed (Chapter II and III). TER2s in the Ler-0 ecotype displayed similar processing as TER2s in Col-0, resulting in a 219nt product lacking the 36nt 3' region (Figure 4-3A, red box).

Strikingly, a different result was obtained with *pot1b-1* mutants. TER2s abundance was unchanged relative to both Ler-0 and Col-0 wild type backgrounds (Figure 4-3B). Using a forward primer aligned to the IS splice junction and a reverse primer aligning to the 3' end (Figure 4-3 A, TER2S + 3'), a robust product was detected. This product was similar in abundance to the TER2s, suggesting that the 3' terminus of TER2s was not cleaved (Figure 4-3B). Notably, the TER2s + 3' product could not be amplified in either Col-0 or Ler-0 wild type backgrounds, suggesting this RNA

intermediate is specific to *pot1b-1* (Figure 4-3B). These data indicate that AtPOT1b influences TER2 metabolism by promoting cleavage of the 3' terminus of TER2 to generate TER2s.

AtPOT1b and TER2 act synergistically to promote chromosome end-protection

Telomerase activity is increased in both the ter2-1 and pot1b-1 mutants. To determine if AtPOT1b and TER2 work in the same genetic pathway, we generated plants doubly deficient for pot1b-1 and ter2-1. Genotyping of these lines has been difficult, but crosses between these two backgrounds result in severe morphological and growth defects consistent with profound telomere disfunction (Surovtseva et al, 2009; Song et al, 2008). The mutants harbored fasciated stems, irregular silique positioning, and continued to produce auxiliary shoots well past the point of senescence of wild type siblings (Figure 4-4A). In wild type, telomere tracts span 2-5kb, but in putative *pot1bter2* mutants, telomeres are dramatically shorter, ranging from 3.5kb to less than 1kb (Figure 4-4B). Because telomeres shorter than 1kb typically elicit a DNA damage response that results in fusions between offending chromosome ends, we performed Telomere Fusion PCR. Abundant telomere fusions were detected in these lines, similar to the ctc1 control (Figure 4-4C). Cytological analysis revealed abundant anaphase bridges, indicative of gross telomere deprotection (Figure 4-4D). In addition, the pot1b/ter2 background exhibited a ~2.5 fold increase in G-overhang signal, similar to *ctc1* or *stn1* mutant backgrounds.

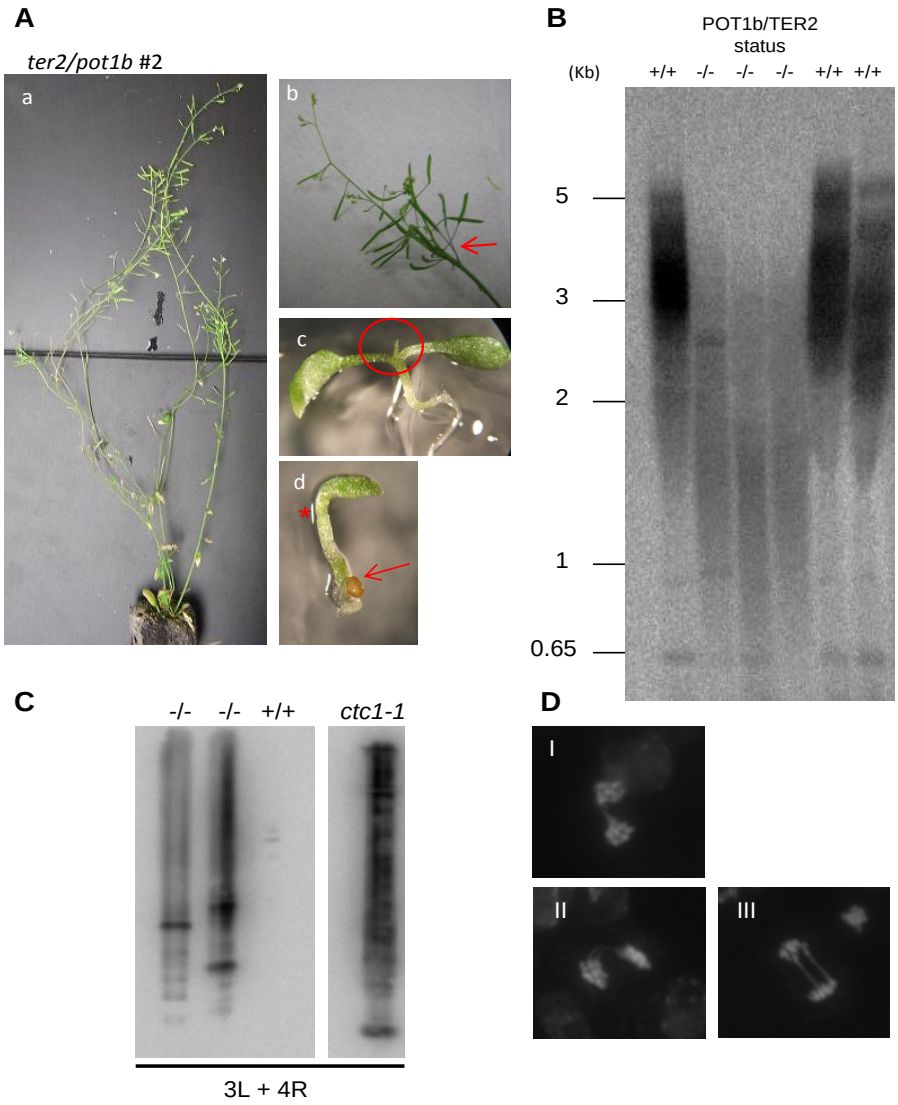


Fig 4-4. POT1b and TER2 cooperate to protect telomeres in *Arabidopsis thaliana*. (**A**) Representative pictures of pot1b/ter2 double mutants (a-b) displaying fasciated stems and irregular silique morphology and placement (b). Second generation seeds displayed reduced germination efficiency (12-33%) and showed growth defects early in development (c-d). (**B**) Bulk telomere analysis by TRF of pot1b/ter2 double mutants and wild type siblings. Presence of either POT1b or TER2 is indicated above each lane. (**C**) Telomere fusion PCR (TF-PCR) on pot1b/ter2 double mutants and wild type siblings. TF-PCR was performed using subtelomeric primers 3L and 4R. PCR products were also cloned and sequenced to confirm the presence of fusions. A *ctc1* mutant was used as a positive control for the reaction. (**D**) Cytology of mitotic chromosomes in *pot1b-2/ter2*-1 (I-III) and wild type (IV). Examples of anaphase bridges are shown.

Discussion

AtPOT1b interacts with TER2 to form a negative regulatory telomerase complex in vivo

Across Eukarya, POT1 proteins are critical for chromosome end-protection and suppression of the DNA damage response. Initial analysis of AtPOT1b over-expression mutants suggested that the ancestral capping function was conserved in *A. thaliana* (Shakirov et al, 2005). However, this model was challenged by the failure to detect telomeric DNA binding for AtPOT1b *in vitro*, and now here, as we show that a null mutation in pot1b-1 does not lead to a substantial telomere uncapping phenotype. We did note a modest decrease in G-overhang signal. Telomerase activity increased in plants lacking pot1b. Remarkably, the fold increase was similar to that observed in a ter2 mutant. This may suggest that POT1b and TER2 regulate activity via similar mechanisms.

Our data indicate that like AtPOT1a, AtPOT1b is involved in telomerase regulation. POT1a and POT1b are physical components of different telomerase RNPs *in vivo* (Cifuentes-Rojas et al, 2011). The TER1 RNP, currently known to consist of POT1a, Dyskerin, and TERT, is critical for telomere length maintenance. Removal of any of these components results in progressive telomere shortening (Surovtseva et al, 2007; Kannan et al, 2008; Fitzgerald et al, 1999, Cifuentes-Rojas et al, 2011). Several accessory factors have been identified for the TER2 RNP, which is surprising given the minimal phenotype ascribed to lines lacking TER2 (Chapter II). TER2 associated factors include Dyskerin, Ku, ATR, and POT1b. This study and other work (Chapter II) has demonstrated a role for both POT1b and TER2 in negatively regulating telomerase

activity. In addition, TER2 levels are elevated in response to DNA damage, corresponding to a decrease in telomerase activity (Chapter II). However, decreased telomerase activity cannot be attributed to all TER2 RNP components. A null mutation within Ku70 does not alter activity levels in Arabidopsis (Nelson A, Shippen D, unpublished data). Therefore, the TER2 RNP may act as a scaffold for critical chromosome-end components.

AtPOT1b interacts with TER2s to form a telomere capping complex

In agreement with the model that TER2 is acting as a scaffold at the chromosome-end we found a severe telomere de-protection phenotype in *ter2pot1b* double mutants. The phenotype we observe is highly reminiscent to that seen in plants deficient for any of the members of the CST complex. Ter2pot1b mutants show reduced fertility, and offspring show decreased germination efficiency. One caveat to for these experiments is that the cross between pot1b and ter2 utilizes two ecotypes, Col-0 and Ler-0, that are genetically distinct. This is most obvious at the TER2 locuse, where Ler-0 contains a TER2 missing the intervening sequence, whereas Col-0 retains the IS (Chapter III). The genotyping assay is more complex, and effort will be required to clarify the genetic status of these plants. However, the telomere phenotype is specific to a cross between these two particular backgrounds, and segregates in a Mendelian fashion.

If the dramatic phenotype observed in the pot1bter2 background is indeed derived from mutations in these two genes, the data suggests that chromosome-end protection is in part mediated by a telomerase RNP complex. This is not without

precedent. Independent TERT or TER mutants in Candida albicans accumulate long G-overhangs, consistent with extensive degradation (Hsu et al, 2007). Furthermore, in budding yeast, telomerase is present in a Ku-dependent manner in the G1 phase of the cell-cycle when telomeres are not being extended. Thus, telomerase association with the telomere may be important for functions outside of DNA replication.

All three POT1 proteins are part of a telomerase RNP complex. This would suggest a redundancy in POT1 proteins in the telomerase regulation. However, the lack of a telomere phenotype in the *pot1b-1* background, but the dramatic phenotype observed in the *pot1bter2* background suggests that there is redundancy at the chromosome end instead. However, the data presented here suggests that shelterin in the Brassicaceae lineage may contain an RNA component. A model for this is presented in Chapter VII.

CHAPTER V

CHARACTERIZATION OF POT1C, A NOVEL MEMBER OF THE ATPOT1 PROTEIN FAMILY WITH ROLES IN TELOMERASE REGULATION AND TELOMERE MAINTENANCE

Summary

Telomeres serve as protective barriers against degradation and end-to-end chromosome fusions. In most organisms, the single strand 3' overhang found at telomeres is protected by telomere-specific OB-fold containing proteins. POT1c is a single OB-fold containing protein derived from a recent partial gene duplication of AtPOT1a. Although POT1c retains extensive nucleotide identity to AtPOT1a, differential splicing gives rise to a protein with <50% amino acid similarity to AtPOT1a. Recombinant POT1c protein demonstrated weak, but non-specific DNA binding *in vitro*. Notably, specific interactions between POT1c and each of the AtTERs were observed. The function of POT1c *in vivo* was explored using RNA interference, which achieved ~70% knockdown (KD). Quantitative RT-PCR and quantitative TRAP of KD plants revealed a ~60% reduction in TER2 levels and a 4-fold reduction in telomerase activity, respectively. These findings argue that POT1c regulates telomerase activity through its interaction with the TER2 RNP. In striking contrast to AtPOT1a and AtPOT1b null mutants, bulk telomeres in POT1c KD lines were dramatically shortened and highly heterogeneous in length. Although no telomere fusions were observed, a G-overhang signal was undetectable in mutant plants. In addition, abundant extra chromosomal

telomeric circles were observed, consistent with recombinational deletion. We conclude that POT1c is an essential component of the telomere capping complex with an additional role in telomerase regulation. Our findings underscore the dramatic evolution of POT1 in the plant kingdom.

Introduction

The conversion from circular to linear chromosomes necessitated a means for replicating the chromosome termini and preventing these ends from being recognized as double strand breaks. This issue was resolved in most eukaryotes by the addition of tandem arrays of G-rich repeats via the telomerase reverse transcriptase. Telomerase is composed of a highly conserved catalytic subunit (TERT) and an integral RNA template (TER), as well as several accessory factors that are essential for proper maturation, activity, and localization of the RNP *in vivo*.

In Arabidopsis, three TER isoforms have been identified, TER1, TER2, and a processed form of TER2, termed TER2s. These TERs form unique RNP complexes *in vivo*, and all three are capable of reconstituting telomerase activity *in vitro*. TER1 is responsible for the canonical telomere length maintenance role, whereas TER2 is responsible for inhibiting telomerase activity during DNA damage. No function has been ascribed to TER2s yet.

Telomeres end in a 3’ G-rich single-strand overhang, termed the G-overhang. The G-overhang is critical for telomere extension, but if unprotected, is a prime substrate for an ATR-mediated DNA damage response (DDR). Consequently, the G-overhang is

carefully guarded by at least two different telomere capping complexes, shelterin, and CST (CTC1/Cdc13, STN1, TEN1) (Price et al, 2010). Shelterin is composed of six proteins (Palm and de Lange, 2008). TRF1 and TRF2 bind directly to double-strand DNA while RAP1 localizes to telomeric DNA through its interaction with TRF2. TIN2 acts as a bridging component between the TRF/RAP1 components and the G-overhang binding heterodimer POT1/TPP1. In most species, POT1 shows high specificity towards ssDNA that is heightened by interactions with TPP1. Loss of POT1 and/or TPP1 is lethal in vertebrates, and results in increased G-overhang signals and activation of a DNA damage response.

The heterotrimer CST complex, comprised of CTC1/Cdc13, Stn1, and Ten1, regulate telomere repeat addition by telomerase, and facilitate the replication of telomeres. In budding yeast, loss of CST components results in abnormally long G-overhangs, elongated telomeres, and activation of a DNA damage response. In vertebrates, the CST complex does not play a major role in telomere length regulation, but is critical for proper G-overhang maintenance and replication of the telomere through interactions with DNA polymerase α primase. Arabidopsis mutants harboring null alleles for any of the CST components display short, deregulated telomeres, massive chromosomal fusions, increased G-overhangs, as well as growth and developmental defects consistent with perturbation of stem cell maintenance (Song et al, 2008; Surovtseva et al, 2009).

Whereas CST in Arabidopsis seems to be well conserved with that found in vertebrates, shelterin has apparently undergone a complex evolutionary transformation.

TPP1, RAP1, and TIN2 sequence homologs cannot be discerned in any plant genome. In contrast, *A. thaliana* encodes 12 TRF-like proteins, of which six show specific telomeric DNA binding via a MYB-extension domain in their C-terminus (Karamysheva, 2004). The POT1 family of proteins has also undergone an expansion. In the Brassicaceae family, two duplication events have occurred, resulting in two full length POT1 proteins, POT1a and POT1b. A third, single OB fold protein, POT1c can be found in the *A. thaliana* genome. POT1 duplication events appear to have coincided with neofunctionalization, as both POT1a and POT1b are components of distinct telomerase RNP complexes. AtPOT1a binds to TER1 and is essential for telomere maintenance and robust telomerase enzymology. Similar to the telomerase accessory factor in yeast, Est1p, POT1a interacts with CTC1 and STN1 *in vitro*. POT1a association with telomeres peaks during S-phase, during which telomerase is thought to extend telomeres. In contrast, POT1b binds both TER2 and TER2s. POT1b negatively regulates telomerase activity, is necessary for proper TER2s maturation, and appears to be part of a telomere capping complex with TER2 (Chapter IV). Here we characterize the third member of the POT1 protein family in Arabidopsis, POT1c. We show that POT1c is a new member of the POT1 family that is confined to *A. thaliana*. POT1c, like POT1b, negatively regulates telomerase, and this function may be mediated through TER2 stabilization. Furthermore, in contrast to POT1a and POT1b mutants, plants deficient in POT1c display profoundly altered telomere architecture indicating that POT1c plays a crucial role in chromosome-end protection.

Materials and Methods

POT1c knockdown, Genetic backgrounds used, Expression analysis, QRT-PCR, and Phylogenetic analysis

~100bp constructs targeting the 5'UTR, Exon 1, Exon 2, and Exon 4 were cloned into the Gateway entry vector pDONR201 using BP clonase (Invitrogen). Clones were sequenced and transferred into the destination vector pB7WGIGW2 using LR clonase II (Invitrogen). These constructs were transformed into Agrobacterium tumefacians (strain GV3101) and subsequently transformed by the floral dip method into *Arabidopsis thaliana* ecotype col. Transformants were selected by growing on Murashige and Skoog basal medium supplemented with 20mg/liter of glufosinate ammonium (BASTA, Crescent Chemicals).. RNA was extracted from tissue for expression analysis using the EZNA RNA extraction kit according to manufacturer's instructions (OmegaBiotek). cDNA was reverse-transcribed by adding 1ug of RNA to 4ul qScript cDNA supermix in a 20ul reaction (Quanta Biosciences). 4ul of a 1:50 dilution of cDNA was added to a reaction containing 300nM of each primer and 10ul of Sso-Advanced SYBR QPCR supermix (Bio-Rad). POT1 transcript levels were tested by QRT-PCR. Fold change was calculated using the ΔΔCT, using GAPDH and TIP41L as internal controls. Primers for these genes, as well as POT1a/b/c, are listed in Supplemental table 1. Phylogenetic analysis of the POT1 proteins.

TRF, PETRA, G-overhang analysis, T-Fusion PCR, Extra Chromosomal Telomere Circle assay

Bulk telomere analysis, PETRA, Native in-gel hybridization, Telomere-Fusion PCR, and ECTA were performed as described previously. DNA for these assays was extracted from total plant tissue using the CTAB method.

Results

Arabidopsis thaliana harbors a third POT1-like protein with a single OB fold

During the initial BLAST of the *A. thaliana* genome for POT1-like proteins, two full-length OB-fold encoding genes were identified, dubbed AtPOT1a and AtPOT1b. In addition, a truncated gene was detected approximately 350 kilobases upstream of AtPOT1a, designated POT1c. Genome sequence analysis revealed that the POT1a and POT1b duplication extends to the base of the Brassicaceae lineage (Beilstein et al, 2012, Figure 1A). In contrast, although POT1c is present in all *A. thaliana* ecotypes sequenced thus far, it is not present in *A. lyrata*, placing the age of the POT1c duplication at ~10mya (Beilstein et al, 2010, Cao et al, 2011), (Figure 5-1A). BLAST analysis of the sequences flanking AtPOT1a and POT1c showed that the 3kb region around POT1c originated from the tandem duplication and inversion of two genes, At2g04395 and At2g04390, from their counterparts At2g05210 and At2g05220. This incomplete duplication removed 1100bp from the 3' end of the new POT1c gene, leaving a region that retains 90% nucleotide identity, with most of the dissimilarity (7%) arising from gaps, and not substitutions. Alternative splicing of POT1c produces two isoforms of

POT1c, one previously annotated at 158aa, and a longer, more prevalent variant of 225 amino acid protein with <50% identity to AtPOT1a (Figure 5-1B, C). POT1c protein is predicted to form a single oligosaccharide/oligonucleotide DNA binding domain (OB-fold, Jpred3 secondary structure prediction; Cole et al, 2008). Typical of other POT1 proteins, AtPOT1a and AtPOT1b contain two OB-folds and a short C-terminal region (Figure 5-1B). The single OB-fold of POT1c is a combination of the first half of the OB1 domain and the second half of the OB2 domain from AtPOT1a, resulting in a chimeric OB-fold (Figure 5-1B, C). Indeed, aside from two gaps of ~100 and 10 amino acids, only 10 residues are different between POT1a and POT1c. Comparison of the POT1c locus between the 511 currently available Arabidopsis genome sequences reveals that 4/10 amino acids in the Col-0 ecotype vary in other *A. thaliana* ecotypes. (Figure 5-1C, red asterisks). Notably, two of the residues within AtPOT1a found to be under positive selective pressure are conserved in POT1c (Figure 5-1C, black asterisks).

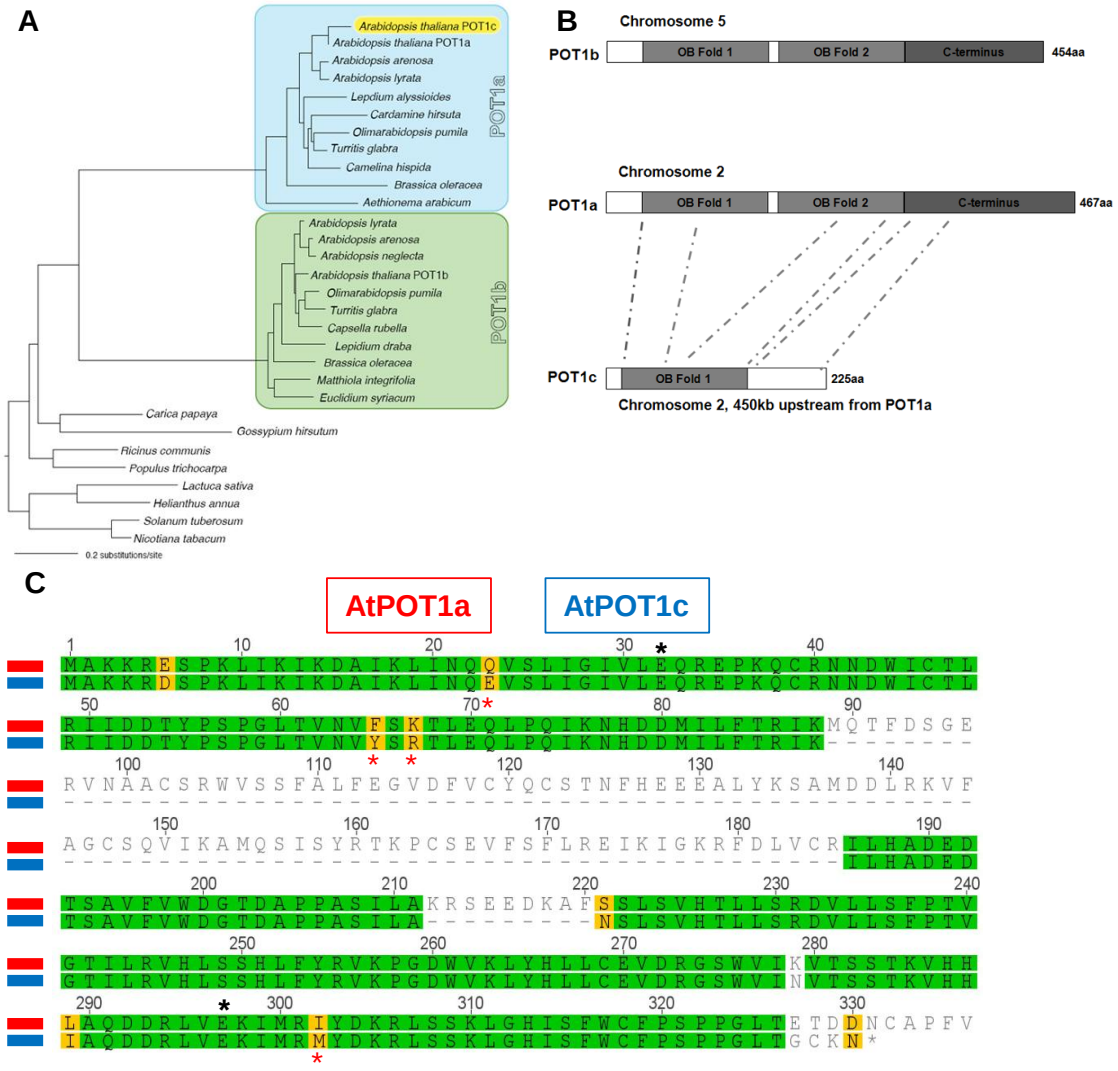


Fig 5-1. AtPOT1c is a recent duplication of AtPOT1a. (**A**) Phylogenetic tree of the POT1 proteinfamily within the Rosid clade showing the relationship between POT1c and the rest of the POT1 family (highlighted). Likelihood tree was produced using the WAG model of amino acid transitions. (**B**) Domain organization within POT1c, based on predicted secondary structure. Secondary structure prediction was performed using Jpred3. (**C**) Alignment of AtPOT1a (red) and AtPOT1c (blue). Residues showing variation within the population of A. thaliana ecotypes for AtPOT1c are indicated by red asterisk. These residues all showed variation between consensus sequence and the corresponding AtPOT1a sequence. All other AtPOT1c variations were completely conserved between ecotypes. Residues within AtPOT1a under positive selection are indicated by black asterisks.

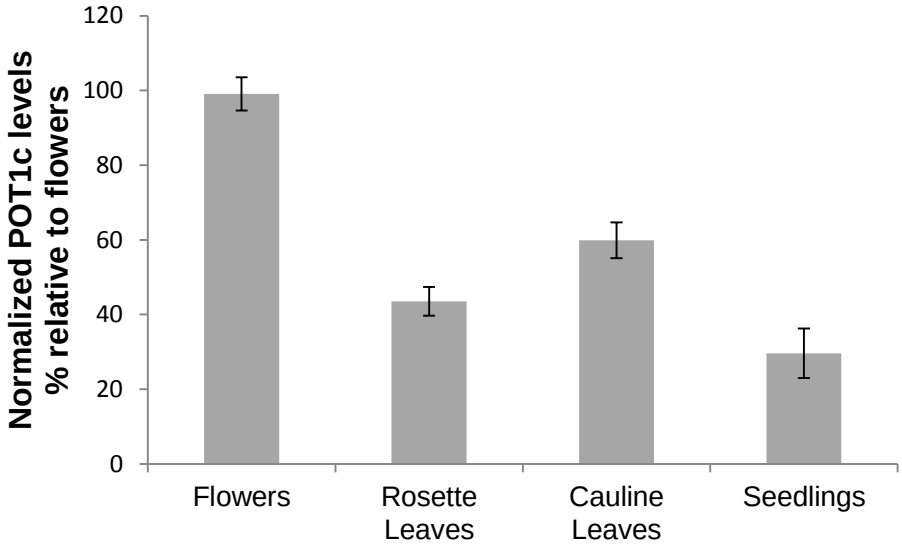


Fig 5-2. AtPOT1c expression levels in *Arabidosis thaliana*. QRT-PCR was performed on n ≥5 independent samples. Normalization was performed against GAPDH. Values were corrected for primer efficiency using LinReg software. Error bars represent S.E.M.

POT1c mRNA, similar to POT1a, is most highly expressed in flowers and cell culture, whereas RNA levels were reduced by ~60% in non-proliferating tissues (leaves and stems) and 80% in developing seedlings (Figure 5-2). On average, POT1c mRNA levels were 3-fold lower in all tissues relative to AtPOT1a.

POT1c interacts with both TER1 and TER2 in vitro and in vivo

Both AtPOT1a and AtPOT1b are telomerase components with specific affinities towards TER1 and TER2, respectively (Cifuentes-Rojas et al, 2011; Chapter II). To determine if POT1c also demonstrated RNA binding, an N-terminal 6xHis-MBP tagged POT1c (HMPOT1c) was expressed in E. coli and purified by nickel resin chromatography (Figure 5-3A). RNA-protein filter binding was performed to test the interaction between HMPOT1c and TER1 and TER2 (Figure 5-3B). Increasing amounts of HMPOT1c, ranging from 0.25ng to 2.5ug was incubated with either TER1, TER2, or the control P4P6 prior to filter binding. His-MBP, purified under similar conditions, was used as a negative, non-specific control (Figure 5-3B, right). Unexpectedly, based on the sequence similarity to POT1a, POT1c bound to both TER1 and TER2. This binding was confirmed in a different manner, co-expressing T7-tagged POT1c and the TERs in rabbit reticulocyte lysate (RRL). Immunoprecipitations were performed, followed by RNA extraction and RT-PCR. Just as in the filter binding assay, POT1c is capable of binding either TER *in vitro* (Figure 5-3C).

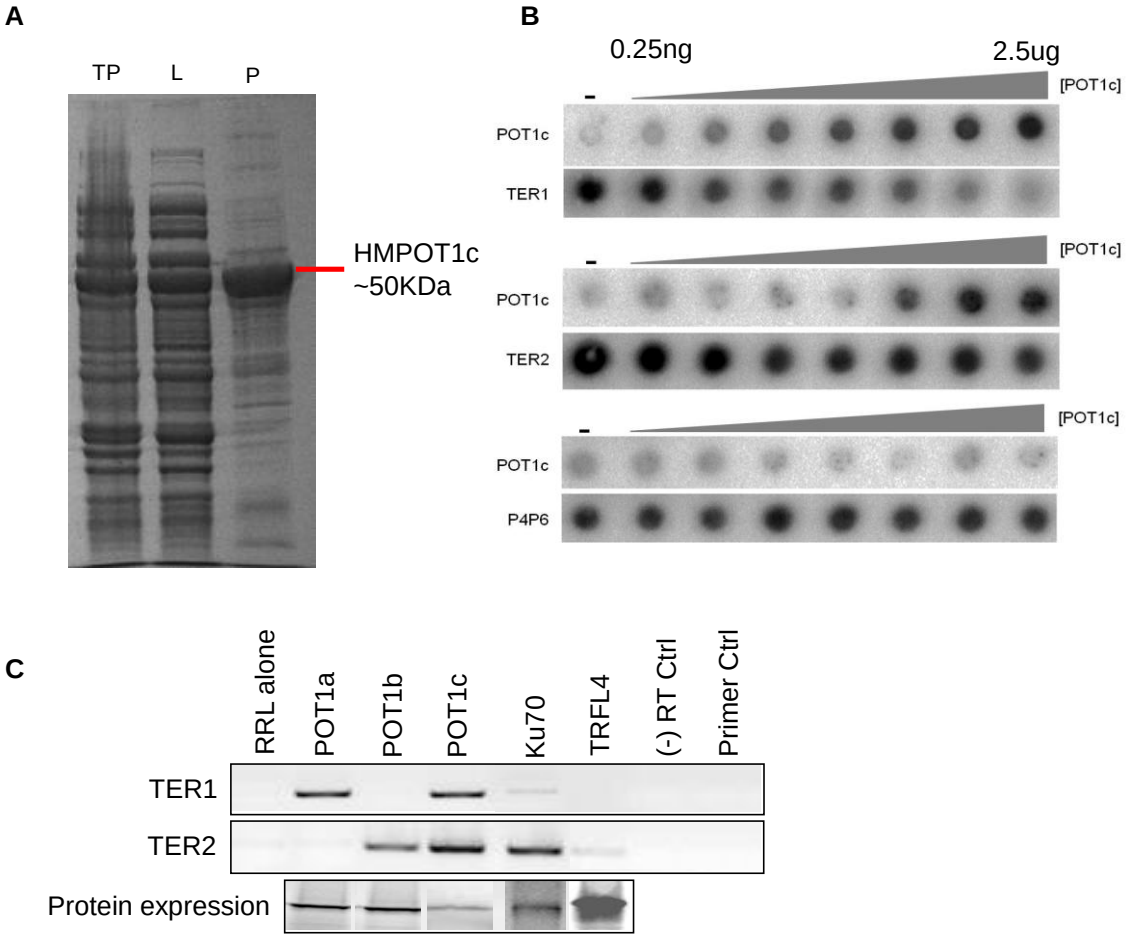


Fig 5-3. AtPOT1c binds both TER1 and TER2 telomerase RNAs. (**A**) Purification of recombinant His-MBP-POT1c using nickel resin. TP= total protein prior to lysis. L= cell lysate prior to purification. P= purified POT1c after elution from nickel resin. (**B**) Filter binding of POT1c with *in vitro* transcribed and 5' end labeled RNA. Increasing concentrations of recombinant POT1c was added to either TER1, TER2, or negative control P4P6. (**C**) Co-IP of T7-POT1c-bound TER1 and TER2. Proteins were co-expressed with the RNA in RRL. Pull downs were performed with αT7-conjugated beads. RNA was extracted, followed by RT-PCR using primers specific to either TER1 or TER2.

POT1c is a negative regulator of telomerase activity

Lines deficient for POT1a show progressive telomere shortening and a decrease in telomerase activity (Surovtseva et al, 2007). Due to sequence similarity between POT1a and POT1c, and the ability of POT1c to bind both TER1 and TER2, we sought to determine if POT1c performed a similar function to POT1a *in vivo*. No null mutant is currently available for POT1c, so RNAi was performed to knockdown POT1c levels. Three different constructs were designed to target different regions of POT1c (Figure 5-4A). Multiple transformations were performed to deliver each of these constructs into a wild type *A. thaliana* Col-0 background. RNA was extracted from flowers of T1 transformants and tested by Q-PCR for decreased levels of POT1c (Figure 5-4B). Some non-specific knock down of POT1a was observed with construct-2 (38% reduction vs 47% for POT1c). However, construct-1 was more specific, reducing levels of POT1c by ~67%, while only reducing POT1a by ~14% (Figure 5-4B). Transformants harboring construct-3 served as a negative control for our experiments, as neither POT1a nor POT1c was knocked down in these lines.

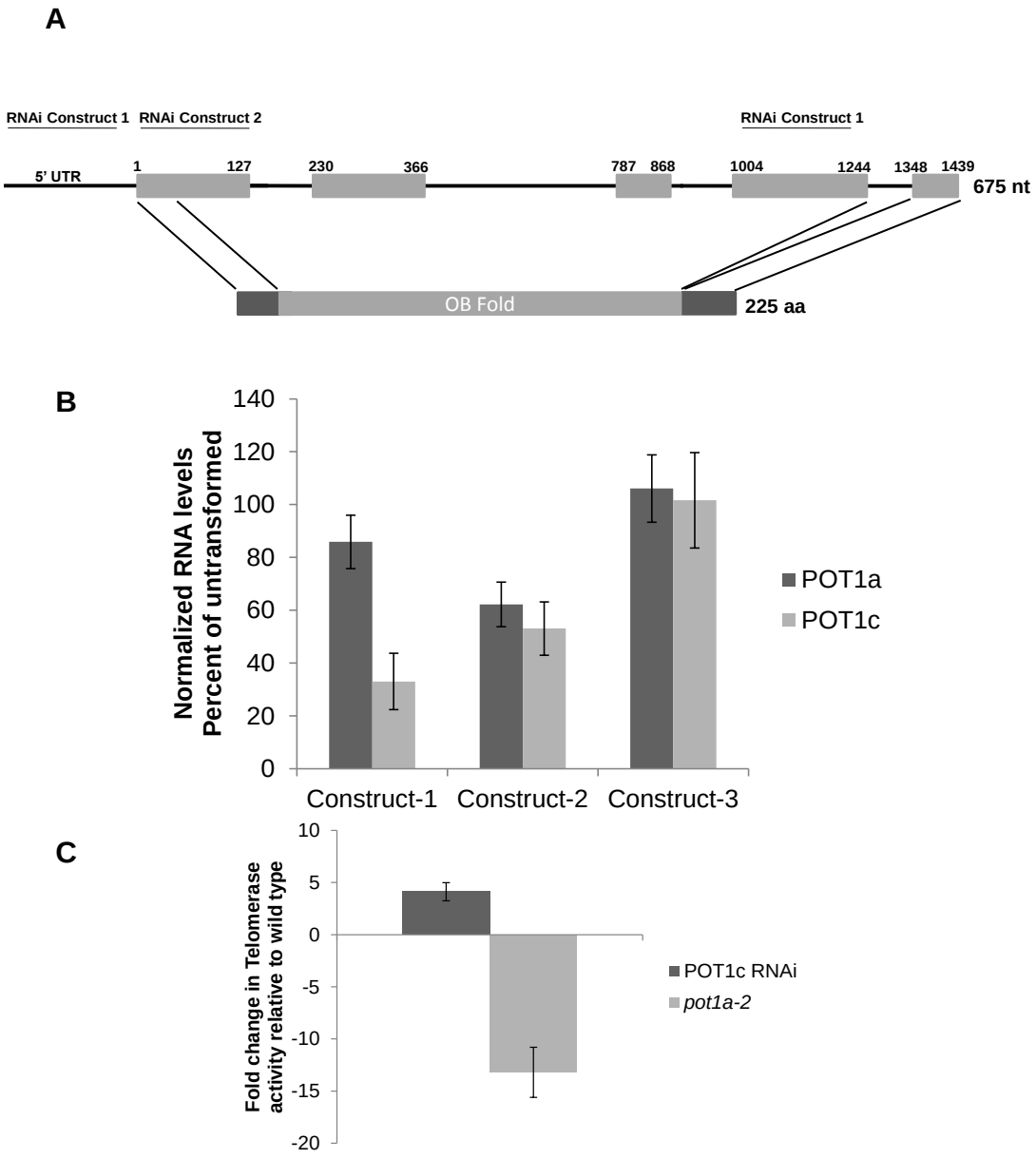


Fig 5-4. Knockdown of POT1c. (**A**) Schematic representation of the POT1c locus. Boxes represent exons, with the first and last nucleotide of the exon marked above. RNAi constructs were designed to the three regions labeled. (**B**) Confirmation of POT1c knockdown by QRT-PCR. QRT-PCR was performed on three independent T1 lines each for the different RNAi constructs. RNA levels were normalized against GAPDH. Primer efficiency was corrected using LinReg. Error bars represent S.D. (**C**) Telomerase activity levels relative in POT1c RNAi lines relative to wild type (construct-3). Values for construct-1 and construct-2 were combined, as they showed a similar increase in telomerase activity. Thus, n=6 for POT1c RNAi, and n=5 for *pot1a-2* null mutants.

We next examined telomerase activity in these knockdown lines by Quantitative TRAP. Total protein was extracted from flower tissue from T1 transformants containing the three RNAi constructs. Contrary to *pot1a* null mutants, which show a ~13-fold decrease in telomerase activity, plants harboring either construct-1 or construct-2 showed an ~4-fold increase in activity compared to construct-3 (Figure 5-4C). These data suggest POT1c and POT1a are regulating telomerase activity differently.

POT1c regulates telomerase activity by stabilizing TER2

TER2 is a potent inhibitor of telomerase activity *in vivo* (Chapter II). TERT shows ~10-fold higher affinity for TER2 than TER1 (Chapter II). Due to this higher affinity, TER2 can sequester TERT into an inactive telomerase complex. This inhibition is most pronounced during rounds of DNA damage, where TER2 becomes the predominant TER isoform, and telomerase activity decreases (Chapter II). In agreement with an inhibitory role for TER2, *ter2* deficient lines exhibit a ~3-fold increase in telomerase activity. Therefore, one possible mechanism by which POT1c could be regulating telomerase activity is through the stabilization of TER2.

To determine if POT1c is stabilizing TER2 *in vivo*, QRT-PCR was performed on RNA from plants harboring the three different POT1c RNAi constructs. Levels of all three TER isoforms were tested from the different RNAi lines. Constructs 1 and 2, but not construct-3, showed a ~50% decrease in both TER2 and TER2s levels (Figure 5-5). Neither POT1c knockdown lines, nor *pot1a-2* showed a significant decrease in TER1 levels (Figure 5-5). Thus, even though POT1c is capable of interacting with both TER1 and TER2 *in vitro*, it regulates telomerase activity *in vivo* through stabilizing TER2.

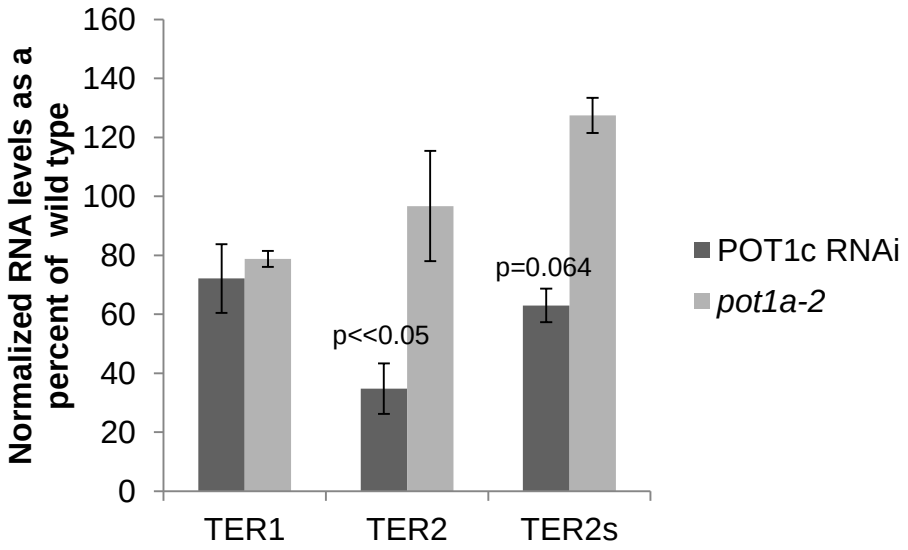


Fig 5-5. TER2 levels are decreased in POT1c RNAi lines. QRT-PCR on the Arabidopsis TER isoforms from three different RNAi lines. RNA levels were normalized to GAPDH. Primer efficiency was corrected for using LinReg. P values were calculated using a two-tailed student's T-test.

POT1c RNAi lines reveal a role in chromosome end-protection

Both POT1a and POT1b mutants display defects in telomere length maintenance (Surovtseva et al, 2007; Shakirov et al, 2005). To determine if knockdown of POT1c perturbs telomeres in Arabidopsis, bulk telomere length was monitored by terminal fragment length analysis (TRF) (Figure 5-6A). Surprisingly, telomeres from the RNAi construct-1 lines were on average 1500bp shorter than the wild type controls (Construct-3) (Figure 5-6A, left). In addition, telomeres were highly heterogeneous, suggestive of the nucleolytic processing seen in plants missing the Arabidopsis CST component CTC1 (Surovtsova et al, 2009). In addition, lines that showed reduced levels of both POT1c and POT1a show a synergistic decrease in telomere length reminiscent of 3^{rd} or 4^{th} generation *pot1a-2* mutants (Figure 5-6A, right). Thus, POT1c and POT1a act via separate mechanisms to regulate telomere length.

POT1c is crucial in G-overhang length regulation and prevention of ECTC

A possible explanation for the highly heterogeneous nature of telomeres in the POT1c RNAi lines is that these telomeres are no longer sufficiently capped. This would expose them to exonucleolytic degradation and potentially engage the DNA damage machinery, leading to chromosomal fusions. To test for uncapping of the telomeres, we first sought to determine the status of the G-overhang in both the RNAi-1 and RNAi-2 KD lines. G-overhangs were analyzed using a native in-gel hybridization technique, using a radio-labeled probe complementary to the G-overhang $(AAATCCC)_3$. Surprisingly, we were unable to detect a signal above background for the POT1c RNAi

lines even though a signal was acquired for wild type and RNAi construct-3 lines. These data suggest that POT1c is important for G-overhang maintenance.

Perturbation in the length of the G-overhang can lead to recognition of the chromosome end as a double strand break, resulting in cell-cycle inhibition and potential chromosome end-to-end fusions. Due to the short nature of Arabidopsis telomeres, and the unique subtelomeric sequences on 8 out of 10 chromosome arms allows for the detection of fusions by a sensitive PCR-based technique, Telomere-Fusion PCR (TF-PCR). We analyzed the POT1c RNAi lines for telomere fusions by TF-PCR and traditional cytological means. However, we were unable to detect either fusions or evidence of anaphase bridges (data not shown).

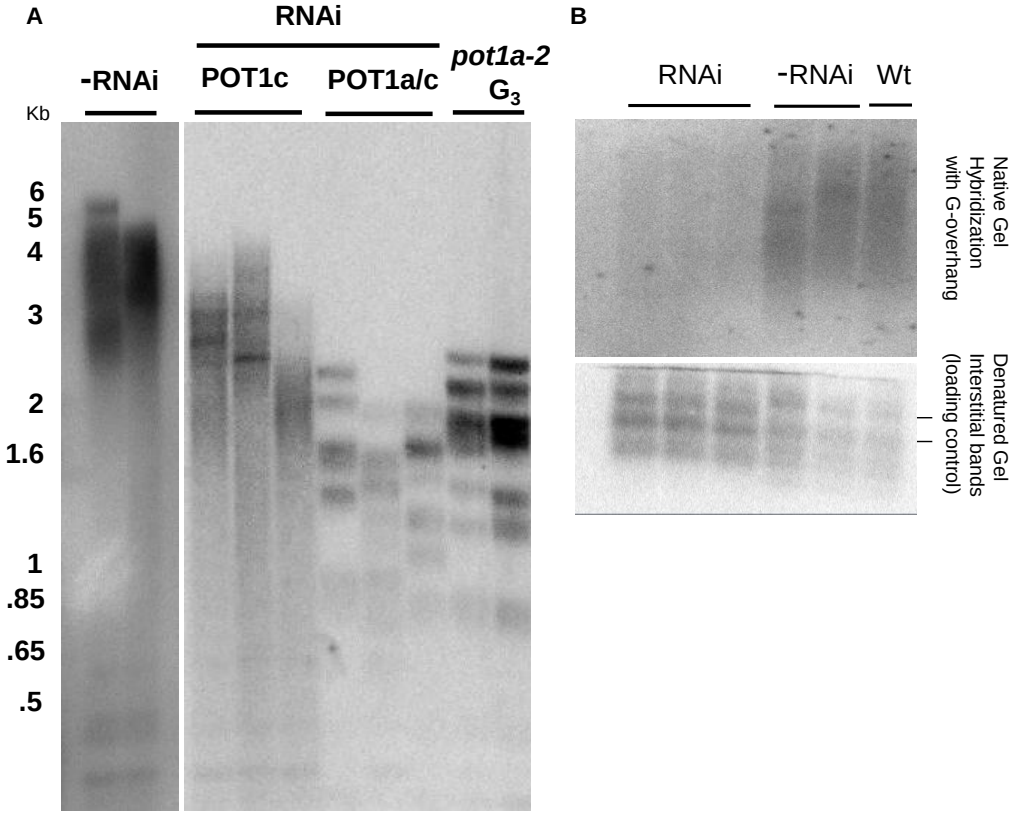


Fig 5-6. Telomere length deregulation in POT1c RNAi lines. (**A**) Terminal Restriction Fragment length (TRF) analysis of bulk telomeres from POT1c RNAi lines. Control lines exhibiting no POT1c RNAi (construct-3) are shown to the left. Ladder is in kilobases. (**B**) Native in-gel hybridization of G-overhangs in POT1c RNAi lines. Loading between samples is standardized using the denatured gel containing the interstitial bands. Quantification was not performed due to the low signal for the RNAi samples.

Since telomeres were short and heterogeneous in the RNAi lines, yet did not seem to be recruited into chromosome fusions, we sought an alternative explanation for the rapid decrease in telomere length in the POT1c RNAi lines. One explanation is telomere rapid deletion (TRD). One hallmark of TRD in Arabidopsis is the formation of telomeric circles (t-circles, Zellinger et al, 2007). We examined the POT1c RNAi lines for t-circle formation using the Extra Chromosomal Telomere Circle Assay (ECTC). This assay utilizes a polymerase with strand displacement activity (phi29) that amplifies long stretches of linear DNA from a circular template. Southern blotting displays a high molecular weight band that migrates close to the top of the gel. At*ku70* mutants have previously been analyzed by the ECTC assay, and act as a positive control for this assay (Figure 5-7). POT1c KD lines showed abundant ECTCs, albeit not as high as the At*ku70* control. The sole loss of POT1a or POT1b does not result in ECTCs, suggesting POT1c has a more direct role in preventing telomere recombination (Figure 5-7).

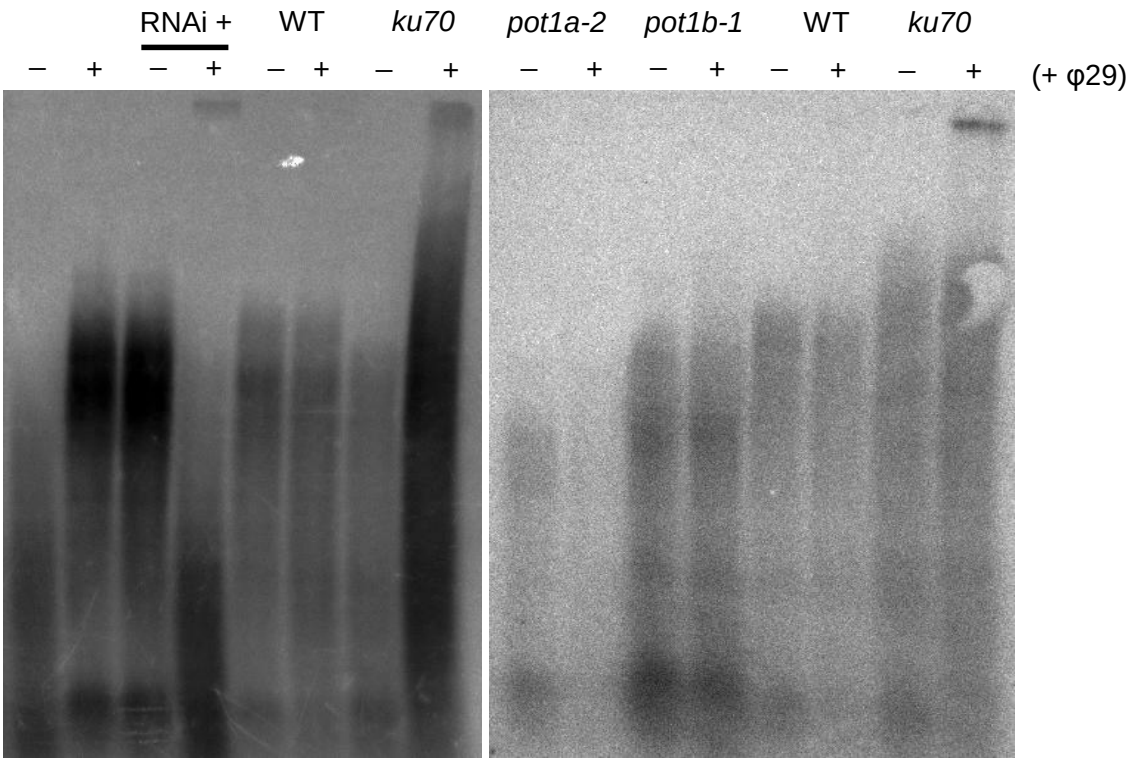


Fig 5-7. POT1c RNAi lines harbor t-circles. An extra chromosomal telomere circle assay (EC-TCA) was performed on POT1c RNAi lines, *pot1a-2* (Surovtseva et al, 2007), *pot1b-1* (Chapter IV), using wild type (WT, Col-0), RNAi construct-3 and *ku70* as controls (Zellinger et al, 2007). + indicates the addition of the polymerase phi29 to the reaction. Phi29 activity on circular DNA generates high MW products that remains at the top of the gel.

Discussion

POT1 duplication has led to an expansion in POT1 functions at telomeres

Tandem gene duplicates are common throughout eukaryotic genomes (Freeling, 2009). Gene retention is less common. Even though the lineage leading up to Arabidopsis has undergone at least two whole genome duplications (WGDs), less than 27% of the Arabidopsis genome harbors duplicated genes (Blanc and Wolfe, 2004). While POT1 is single copy in many organisms, including the moss Physcomitrella, the data presented here indicate that POT1 was subjected to three independent duplication events in angiosperms. One of these occurred in grasses, as maize is reported to contain two POT1 proteins, one of which binds DNA (Shakirov et al, 2009). The second event originated from a WGD that can be traced to the base of the order Brassicales, resulting in two POT1 genes in the Brassicacea family (Beilstein et al, in prep). The last duplication event is specific to *A. thaliana*, resulting in a third POT1 gene present in all *A. thaliana* ecotypes. POT1 duplication is not confined to the plant kingdom. *C. elegans* carries four single OB-fold POT1-like proteins, whereas mouse and tetrahymena both contain two functional POT1 proteins. Thus, a redundancy in telomere specific OB-fold containing proteins is not detrimental to a genome per se.

POT1c, like AtPOT1b, is a negative regulator of telomerase activity

While it is still too early, evolutionarily speaking, to know if POT1c is undergoing neo or sub-functionalization, our data indicate POT1c is acting in concert, not with its parent, POT1a, but with the POT1a ortholog, POT1b. While POT1c can bind to both TER1 and TER2 *in vitro*, it regulates telomerase activity via the TER2 RNP *in*

vivo. This regulation is most likely through stabilization of TER2, as POT1c KD lines showed a marked decrease in TER2 and TER2s, but not TER1. It is not yet clear how POT1c stabilizes TER2. POT1c RNAi did not result in a noticeable shift in TER2 isoforms to indicate processing was inhibited. Instead, overall TER2 levels declined. Thus is is currently unclear if POT1c is stabilizing both TER2 and TER2s, or only TER2.

The current model for TER2 is that it inhibits telomerase due to the higher affinity TERT has for TER2 over TER1. TER2 levels are increased in response to DNA damage concomitantly with a decrease in telomerae activity. However, complete loss of TER2, POT1b, or knockdown of POT1c results in an increase in telomerase activity independent of a DNA damage response. The large number of molecules devoted to inhibiting telomerase suggests a complex need for telomerase regulation during normal telomere maintenance.

POT1c is important for chromosome end-protection in Arabidopsis

Single OB-fold proteins typically oligomerize, thus allowing broader substrate recognition and higher affinity (Theobald et al, 2003; Kerr et al, 2003). OB-fold containing proteins often use these domains in a modular fashion, with one OB-fold providing higher substrate affinity to another OB-fold. In the case of RPA, the OB-folds do not need to be contained within the same protein, but can lie within an entirely different subunit. Both of the G-overhang binding complexes in vertebrates, CST and POT1/TPP1, recognize DNA through multiple OB-folds. TPP1 enables POT1 to bind tighter to ss telomeric DNA, whereas telomere specific binding requires all three CST

components. Preliminary evidence indicate POT1c is capable of homo-oligomerizing *in vitro,* but the multimeric state of POT1c *in vivo* remains to be elucidated.

Although POT1c consists of a single OB-fold, mutants display a broad array of severe defects associated with chromosome-end protection. This is in striking contrast to *pot1a* or *pot1b* backgrounds. Remarkably, the heterogeneous telomeres, minimal G-overhang signal, and abundant T-circles associated with depletion of POT1c did not trigger telomere fusions. This suggests that other factors are still present to repress the DNA damage response. POT1c may not be sufficient by itself for chromosome-end protection, but instead is necessary for the recruitment or stabilization of a larger telomere capping complex. Similar to POT1a, POT1c interacts with CTC1 *in vitro* (Data not shown). Whether this interaction is physiologically significant is not clear. One model would predict that POT1c stabilizes CST at telomeres outside of S-phase, and a transition to POT1a occurs to facilitate telomere extension by telomerase. Biochemical data to support this argument are lacking. However, the phenotype observed in POT1c knockdown lines hint at an evolutionarily conserved role in chromosome end-protection.

A novel means of carrying on the family name

Due to its similarity to POT1a, we originally predicted that POT1c would perform a similar function at telomeres in Arabidopsis. However, POT1c RNAi lines generated a telomere-phenotype not seen in *pot1a-2*. The question then arises of how POT1c has acquired this role, and what is its significance. The answer may lie with the CST complex, since there is evidence of a POT1c-CTC1 interaction most likely retained from the ancestral POT1a. Intriguingly, the phenotype from POT1c knockdown lines

closely mimics ctc1 mutants, with the exception of the lack of telomere fusions. Thus, the possibility remains for an alternative telomere capping complex in *Arabidopsis thaliana* containing a member of the POT1 family of proteins.

CHAPTER VI

PARAMETERS AFFECTING TELOMERE-MEDIATED CHROMOSOMAL TRUNCATION IN *ARABIDOPSIS THALIANA**

Summary

Conversion of a double-strand break (DSB) into a telomere is a dangerous, potentially lethal event. However, little is known about the mechanism and control of de novo telomere formation (DNTF). DNTF can be instigated by the insertion of a telomere repeat array (TRA) into the host genome, which seeds the formation of a new telomere resulting in chromosome truncation. Such events are rare and concentrated at chromosome ends. Here we introduce tetraploid *Arabidopsis thaliana* as a robust genetic model for DNTF. Transformation of a 2.6kb TRA into 4X plants resulted in a DNTF efficiency of 56%, five-fold higher than in 2X plants and 50-fold higher than in human cells. DNTF events were recovered across the entire genome, indicating that genetic redundancy facilitates recovery of DNTF events. Although TRAs as short as 100bp seeded new telomeres, these tracts were unstable unless they were extended above a 1kb size threshold. Unexpectedly, DNTF efficiency increased in plants lacking telomerase, while DNTF rates were lower in plants null for Ku70 or Lig4, components of non-

---

*Reprinted with permission from Nelson A.D.L., Lamb J.C., Kobrossly P.S., and Shippen D.E. 2011. Parameters affecting telomere-mediated chromosomal truncation in *Arabidopsis thaliana*, *Plant Cell* **23**, 2263-2272.

homologous end-joining (NHEJ) repair pathway.We conclude that multiple competing pathways modulate DNTF, and that tetraploid *Arabidopsis* will be a powerful model for elucidating the molecular details of these processes.

Introduction

The natural ends of chromosomes are distinguished from double-strand DNA breaks (DSBs) because they are packaged into telomeres, specialized nucleoprotein complexes assembled on a terminal array of short DNA repeats (TTTAGGG in *Arabidopsis* and most plants). The telomere repeat array (TRA) is mostly comprised of double-strand DNA, but terminates in a short G-rich single-strand 3' protrusion termed the G-overhang (Verdun and Karlseder 2007). Telomere length is maintained in a dynamic range by opposing processes that shorten or extend the TRA (Shore and Bianchi 2009). Incomplete DNA replication, nucleolytic degradation and recombination events cause telomeres to shorten (Crabbe et al. 2004; Ferreira et al. 2004), while telomerase extends the TRA (Collins 2006). In vertebrates the TRA is protected by a six-member protein complex termed shelterin (de Lange 2005), however in *Arabidopsis* and budding yeast, the predominant end protection complex is CST (Cdc13/CTC1, Stn1, and Ten1) (Bertuch and Lundblad 2006; Surovtseva et al. 2009). Loss of core telomere capping components is highly deleterious, triggering DNA damage checkpoints, chromosome end-joining reactions and widespread genome instability (Baumann and Cech 2001; Nugent et al. 1996; Puglisi et al. 2008; Surovtseva et al. 2009; van Steensel et al. 1998).

Chromosomal DSBs are typically resolved by non-homologous end-joining (NHEJ) and homologous recombination (HR). They can also be subjected to "chromosome healing" in which telomerase adds telomere repeats at the break site to establish a new telomere (Murnane 2010; Pennaneach et al. 2006). This latter process protects the nascent terminus from subsequent repair activities, but *de novo* telomere formation (DNTF) is perilous as it leads to deletion of the acentric distal chromosome fragment. Terminal deletions and DNTF are associated with several genetic disorders including α-thalassaemia and some forms of mental retardation (Flint et al. 1994; Wilkie et al. 1990) as well as cancer (Lee and Myung 2009).

DNTF is best understood in *Saccharomyces cerevisiae.* In this setting, DNTF requires the telomere capping protein Cdc13 and the telomerase-accessory factor Est1, which collaborate in the recruitment and/or activation of telomerase following resection of the 5' end of the TRA by the MRX (Mre11/Rad50/Xrs2) nuclease (Diede and Gottschling 2001; Larrivee et al. 2004). In the absence of a TRA, telomerase is essential for the establishment of a new telomere (Pennaneach and Kolodner 2004). However, if the break occurs adjacent to a TRA, telomerase is dispensable and DNTF relies on double-strand telomere binding proteins which presumably assist with the formation of a protective cap on the new telomere (Negrini et al. 2007). The KU70/80 heterodimer also plays an important, but enigmatic role in yeast DNTF. Although Ku is an essential component of the NHEJ DSB repair pathway, it is also required to protect telomeres from end-joining reactions (DuBois et al. 2002). In addition, Ku binds the telomerase

RNA subunit, and this association is essential for DNTF at sites lacking a TRA (Stellwagen et al. 2003).

DNTF has been studied in vertebrates (Bae and Baumann 2007; Barnett et al. 1993; Hanish et al. 1994), but much less is known about the process. The DNTF assay involves transgenic introduction of a TRA, which results in chromosome truncation when the non-telomeric region of the construct integrates into an internal region of the chromosome and the TRA is recognized and established as a bona-fide telomere. DNTF is supported by a TRA as short as 250bp, and requires the vertebrate telomere repeat sequence (TTAGGG) (Hanish et al. 1994; Okabe et al. 2000). DNTF is promoted by a double-strand telomeric DNA binding component of shelterin component, TRF1. As in yeast, conversion of the TRA into a functional telomere does not require telomerase (Gao et al. 2008; Okabe et al. 2000). Finally, DNTF events are rare and are preferentially recovered near endogenous telomeres (Diede and Gottschling 1999; Fortin et al. 2009; Gao et al. 2008), presumably reflecting the aneuploidy associated with chromosome truncation.

The ability of transgenic TRAs to acquire telomere function in plants was first demonstrated by *Agrobacterium*-mediated transformation of immature embryos from *Zea mays* (Yu et al. 2006). In subsequent work, maize chromosomes were truncated using TRAs to create plant minichromosomes (Yu et al. 2007). While the frequency of DNTF events was low, a whole arm truncation was recovered in a spontaneous tetraploid event (Yu et al. 2007). This finding suggests that a major barrier to studying DNTF could be overcome by genetically buffering chromosome truncation.

Synthetic *Arabidopsis thaliana* tetraploids have served as models for understanding the effects of ploidy on plant development for some 40 years (Santos et al. 2003). Notably, large deletions that are lethal in the gametophytic generation of diploid *Arabidopsis* can be propagated in tetraploid (4X) plants (Vizir and Mulligan 1999). Here we introduce tetraploid *Arabidopsis* as a robust new model for DNTF. We observe a remarkably high rate of telomere truncation events, occurring throughout the entire *Arabidopsis* genome. In addition to testing the effects of cis-acting sequence requirements for DNTF, we exploit the genetic tractability of *Arabidopsis* to explore the contribution of telomerase and NHEJ components on DNTF. Unexpectedly, we find that although telomerase is required to sustain new telomeres over the long-term, it is inhibitory to DNTF. In contrast, both Ku and LIG4 are essential for efficient DNTF. These findings argue that multiple competing pathways influence the formation of new telomeres in higher plants.

Materials and Methods

Arabidopsis lines, tetraploidization, and transformation

The *lig4*, *tert*, and *ku70* T-DNA insertion lines have been previously described (Heacock et al. 2007; Riha et al. 2001; Riha et al. 2002). 4X mutant lines were generated by application of 2 microliters of a 0.1% colchicine solution to the apical meristem of 7-14d seedlings (second generation homozygous (G2) for the various T-DNA insertion). Seeds were collected from large flowers. Cytological confirmation of tetraploidy was

performed by counting chromomeres of DAPI stained nuclei (Yu et al. 2006). G3 (4X) plants were then transformed by the floral dip method (Zhang et al. 2006).

Plasmid construction

The 2.6kb, 950bp, 800bp, 700bp, 400bp, and 200bp TRA derivatives of pWY86 were obtained by transforming pWY86 into Stbl2 cells (Invitrogen) and screening individual colonies by restriction digestion with *Eco*RI and *Bgl*II. A 100bp TRA, 950bp TRA, jumbled TRA, UAS, vertebrate TRA, and inverted TRA were constructed using a modified PCR reaction described in detail in SI and cloned into a Gateway entry vector. These constructs were then transferred to *Agrobacterium* destination vectors pBGW or pBWG by the LR clonase reaction (Karimi et al. 2002).

TAIL-PCR

To determine sites of T-DNA insertion or DNTF, TAIL-PCR was performed using both the mTAIL and hiTAIL methods (Liu and Chen 2007; Sessions et al. 2002). Primer sequences are shown in SI. TAIL-PCR products were cloned into the TOPO-TA 2.1 vector and sequenced using the M13 forward primer (Invitrogen).

PETRA and telomere fusion PCR

PETRA and TF-PCR was performed as described in (Heacock et al. 2004) with the following modifications. The PETRA-T reaction was followed by multiple PETRA-A reactions using the PETRA-A primer and either of two primers specific to pWY86: PWY86#1 and pWY86#3 or either of two primers specific to pBGW: PBGW#3 and PBGW#4 (See SI for primer sequences).

Results

Efficient production and recovery of DNTF events in tetraploid *Arabidopsis*

To investigate whether tetraploid *A. thaliana* could serve as a model for DNTF, we generated tetraploids by treating the meristem of 7-14d seedlings with a 0.1% colchicine solution (Henry et al. 2005). Potential tetraploids were initially screened by examining flower size (Figure 6-1A), which increases with ploidy (Henry et al. 2005). Plants grown from seeds of enlarged flowers were then subjected to a cytological screen, which revealed the presence of approximately 18 chromomeres (Figure 6-1B) (Yu et al. 2006). While it was not possible to definitively determine chromosome number in this assay, the data are consistent with tetraploidy.

A TRA-bearing construct previously used to seed DTNF in maize, pWY86 (Yu et al. 2006), was transformed into wild type diploid (2X) and tetraploid (4X) *Arabidopsis* (Col ecotype) (Figure 6-1C). For each construct approximately 100 transgenic lines containing pWY86 (in some cases from two independent transformation events) were selected for basta resistance and then screened for DNTF using a modification of the PCR assay, Primer Extension Telomere Rapid Amplification (PETRA) (Heacock et al. 2004) (Figure 6-1C). In PETRA, the forward primer anneals to the G-overhang, a critical feature of a functional telomere, while the reverse primer binds a unique subtelomeric sequence. To detect DTNF, two PETRA reactions were performed with reverse primers that bind adjacent to each other 46bp (PWY86-1) or 350bp (PWY86-3) upstream of the TRA in the pWY86 construct (Figure 6-1C). Staggered products are generated when T-DNA insertion leads to DNTF rather than integration at an internal site in the

chromosome (Figure 6-1C and Figure 6-2A). The endogenous telomere on the left arm of chromosome 3 (3L) was monitored as a control.

PETRA products were obtained from two independent transformation experiments using 4X wild type *Arabidopsis* (T1=35/60 (58%), T2=26/48 (54%)) (Figure 6-2D) for an average DNTF frequency of 56%. The congruence of these two data points and other results in duplicate transformation experiments (Table 6-1) demonstrated that the DNTF assay was reproducible. In contrast to results with 4X plants, only 10% of the transgenic events characterized in 2X plants led to DNTF (12/126). The elevated incidence of DNTF in 4X plants did not reflect an increased frequency of T-DNA insertion. Southern blot analysis using a probe against the BAR gene (see Figure 6-1C) showed the same number of T-DNA insertions in 2X and 4X transformants (Figure 6-3B).

Consistent with telomerase action at or near the chromosome break site, PETRA products were heterogeneous in length (Figure 6-2A), resembling those generated from the endogenous 3L telomere (Figure 6-2B). Furthermore, the nascent telomeres acquired a G-overhang, as pre-treatment of the reaction with T4 DNA polymerase, which harbors 3'-5' exonuclease activity, blocked PETRA product synthesis (Figure 6-2C). Finally, Southern blot analysis of lines bearing potential truncation events confirmed DNTF (Fig 6-3B).

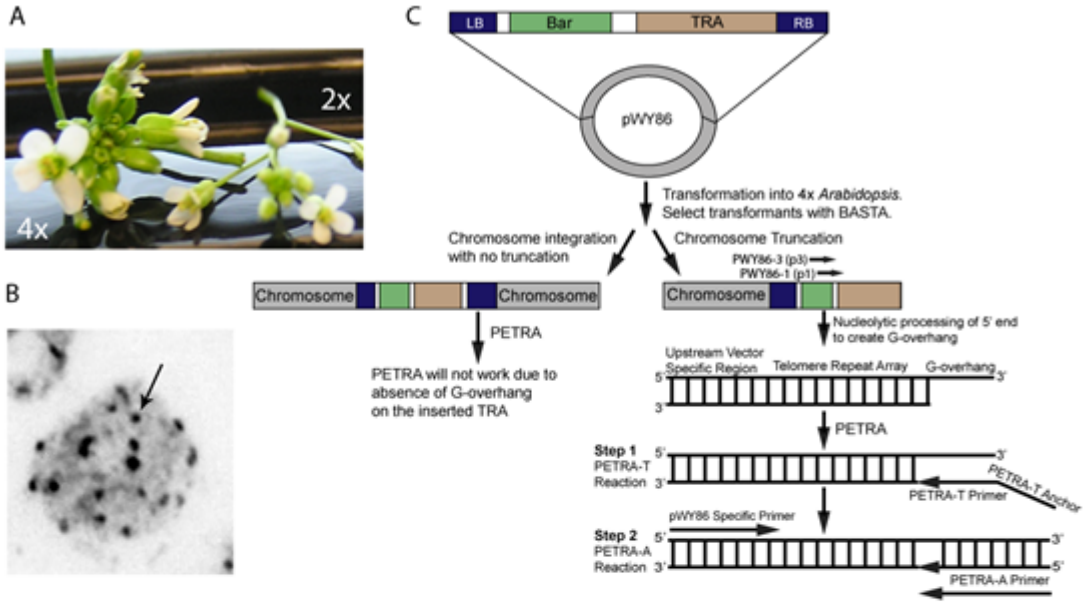


Fig 6-1. The Arabidopsis DNTF system.
(**A**) Increased flower size in tetraploid Arabidopsis. Flowers from tetraploid (4X) and diploid (2X) Arabidopsis are shown. (**B**) Light microscopy of 4X nuclei. Ploidy was confirmed by counting chromomeres (arrow), the heterochromatic region near centromeres. In this image, 18 distinct chromomeres are visible. (**C**) Schematic diagram depicting the pWY86 vector system and the PETRA assay. A T-DNA construct containing the BAR gene (for BASTA selection), LB (left border), RB (right border), and a telomere repeat array (TRA) in pWY86 is transformed into 4X Arabidopsis. The T-DNA can be fully integrated (left) or result in chromosome truncation and DNTF (right). Transformants are screened by PETRA using PWY86-1 (p1) or PWY86-3 (p3) and the PETRA-T primer, which will amplify products with an accessible G-overhang, reflecting DNTF.

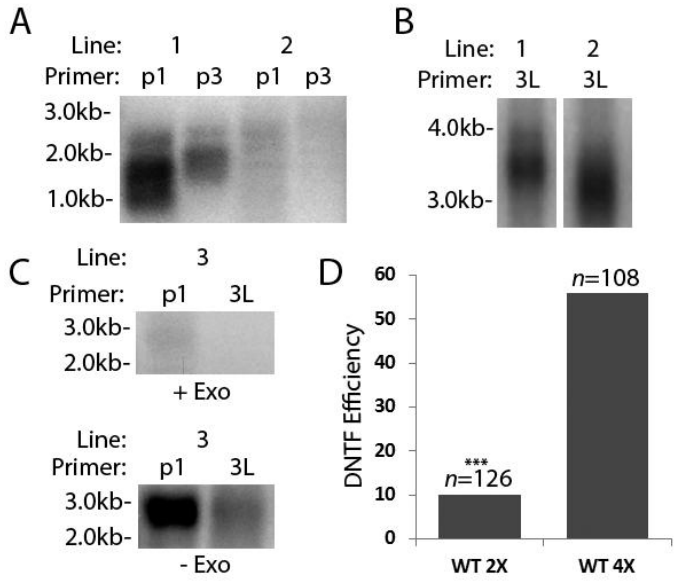


Figure 6-2: Detection of DNTF events by PETRA.
(**A**) PETRA results for 4X wild type carrying a 2.6kb transgenic TRA. Two PETRA reactions with primers p1 or p3 (Fig. 1C) were performed for each plant transformed with pWY86. Line 1 shows the expected staggered size products. Line 2 shows non-specific background bands. Molecular weight markers are indicated. (**B**) PETRA results for the endogenous 3L telomere from two transformants. (**C**) PETRA products are dependent on the presence of a 3' overhang. PETRA was performed on transformants treated with T4 DNA polymerase (+ exo) for 30 min (top) or untreated (- exo; bottom). Results using either pWY86 primer (1) or a primer specific to the endogenous 3L telomere (e3L) are shown. (**D**) Comparison of DNTF efficiency in 2X versus 4X Arabidopsis. Efficiency was calculated by dividing the number of DNTF events by the total number of lines screened (n). *** indicates a statistically significant difference (p-value > 0.0001) in the DNTF rate compared to 4X wild type transformed with a 2.6kb TRA.

Table 6-1: DNTF efficiencies in different genetic backgrounds.

Background	Vector	TRA Length	% DNTF
WT 2X	pWY86	2600	10% (12/126)
WT 4X	pBGW	100	16% (13/80)
WT 4X	pBGW	200	25% (24/96)
WT 4X	pWY86	400	35% (28/80)
WT 4X	pWY86	700	37.5% (36/96)
WT 4X	pWY86	800	47% (47/100)
WT 4X	pWY86	900	53.5 (68/127)
WT 4X	pBGW	950	55% (53/96)
WT 4X	pWY86	2600	58% (35/60) 54% (26/48)
ku70	pWY86	2600	2% (2/88)
lig4	pWY86	2600	35% (26/73)
G4 *tert*	pWY86	2600	71% (68/96) 74% (53/72)
WT 4X	pBGW	750 (TTAGGG)	19% (15/80)
WT 4X	pBGW	500 (TGGTTGAT)	1.3% (1/80)
WT 4X	pBGW	600 (UAS)	0% (0/100)
WT 4X	pBGW	850 (GGGATTT)	0% (0/80)

A description of lines and constructs used, along with the DNTF frequency recovered from these lines.

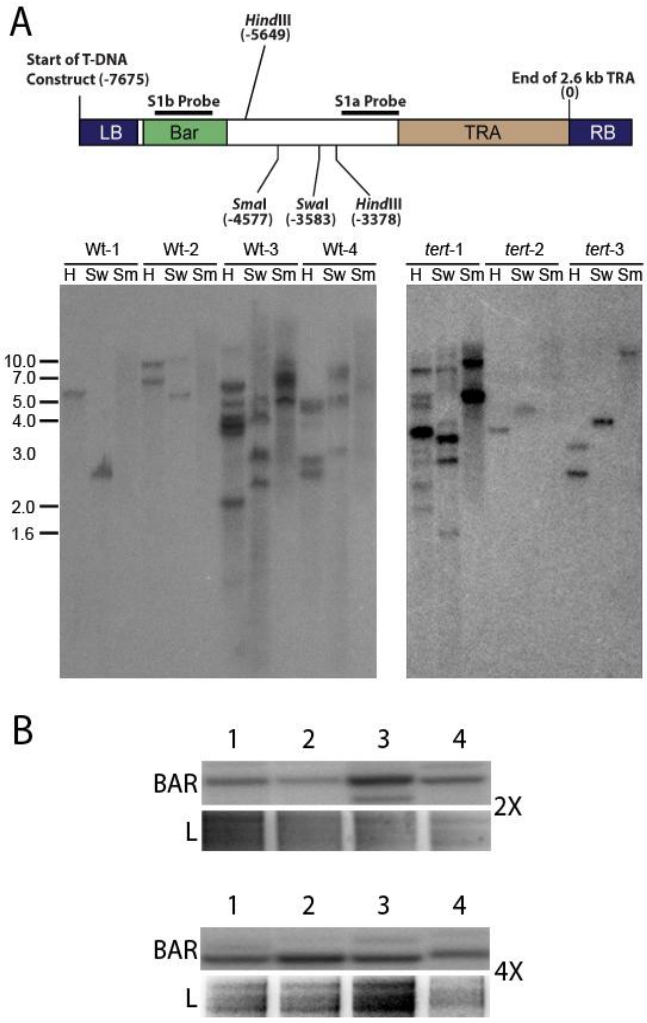
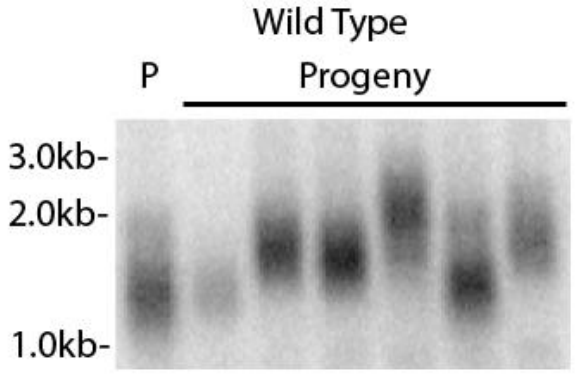


Figure 6-3. DNTF detected by Southern blot analysis.
(**A**) Detailed schematic diagram of PWY86 (top). Restriction sites for HindIII (H), SwaI (Sw), and SmaI (Sm) are shown along with nucleotide positions. Southern blot analysis of A. thaliana transformants (bottom). Results for 4X tert and 4X wild type plants transformed with the 2.6kb TRA using the probe indicated in the top diagram are shown. PETRA revealed DNTF in Wt-3, Wt-4, tert-1, tert-2, and tert-3 transformants, but not in Wt-1 or Wt-2. Hybridization products from DNA digested with HindIII are expected to be ~500bp smaller than SwaI products and 1.5kb smaller than SmaI products. Molecular weight markers are in kilobase pairs. (**B**) Southern blot analysis of 2X and 4X wild type lines carrying the 2.6kb transgenic TRA. The blot was hybridized using a probe for the Basta resistance gene (BAR). L (loading) shows EtBr stained gels prior to transfer.

Telomeres in *A. thaliana* (Col ecotype) typically span 2-5kb (Shakirov and Shippen 2004). To investigate whether new telomeres were subjected to the same length regulation as endogenous telomeres, PETRA was performed in the next plant generation (Figure 6-4A). 4X wild type lines transformed with a 950bp TRA showed on average a 650bp increase in telomere length upon DNTF (Table 6-2). These telomeres were elongated further in the next generation (T2) (Figure 6-4A and 6-4B). Telomeres generally increased in size, reaching up to 3.4kb in T2 (Figure 6-4B). Conversely, new telomeres established using a 2.6kb TRA seed, which falls within the wild type range of endogenous *Arabidopsis* telomeres, did not increase in length (Table 6-2). Thus, the *de novo* telomeres formed in 4X *Arabidopsis* mimic endogenous telomeres in their architecture and length regulation.

A



B

T1 Parent		T2 Offspring	
	Length (bp)	Avg Length (bp)	Avg Increase (bp)
1	1525	1658 +/- 207	133
2	1050	1255 +/- 135	205
3	2300	2908 +/- 80	608
4	2800	3408 +/- 350	608
5	2750	3240 +/- 241	490

Figure 6-4. New telomeres function like native telomeres. (**A**) Parent-progeny analysis of new telomeres. PETRA results are shown for first (T1) and second (T2) generation 4X wild type transformants carrying a 950bp TRA. P denotes parent (T1). Results for six progeny (T2) lines are shown. (**B**) Compilation of PETRA product lengths in T1 4X wild type parents and their T2 progeny transformed with the 950bp TRA. Progeny from five parent lines were analyzed by PETRA. Six progeny were analyzed per parent.

Table 6-2. Length analysis of recovered TRAs

Initial TRA (bp)	TRAs recovered	
	Range (bp)	Avg. (bp)
2600	950-4100	2562 +/- 574
950	850-2925	1608 +/- 359
800	850-3200	1458 +/- 423
700	650-2950	1361 +/- 334
400	825-1912	1436 +/- 366
200	900-2900	1732 +/- 243
100	650-1900	1142 +/- 251

The size of recovered PETRA products in DNTF lines was determined for each length of TRA used in our analysis.

DNTF events occur at random sites in the Arabidopsis genome

We mapped the sites of DNTF in 4X and 2X *Arabidopsis* using thermal asymmetric interlaced (TAIL) PCR (Sessions et al. 2002). We randomly selected ten diploid and thirty tetraploid DNTF lines, and obtained PCR products for 9/10 2X lines and 25/30 4X lines. In all the 2X lines, a new telomere formed close to the endogenous chromosome end (Figure 6-5, blue bars). In contrast, the 25 chromosome truncation events mapped in 4X plants were widely distributed throughout the entire genome (Figure 6-5, grey bars), consistent with random integration of T-DNA (Kim et al. 2007). Strikingly, several DNTF events resulted in large deletions, including the loss of 20Mb from the right arm of chromosome 1 and a truncation event within the centromere-flanking region of chromosome 4 (Figure 6-5, arrowheads) (Kumekawa et al. 2001). Whether this centromere truncation affects chromosome segregation is unknown. The elevated incidence of DNTF and recovery of large chromosomal deletions indicates that tetraploid *Arabidopsis* can be exploited as a robust model for DNTF.

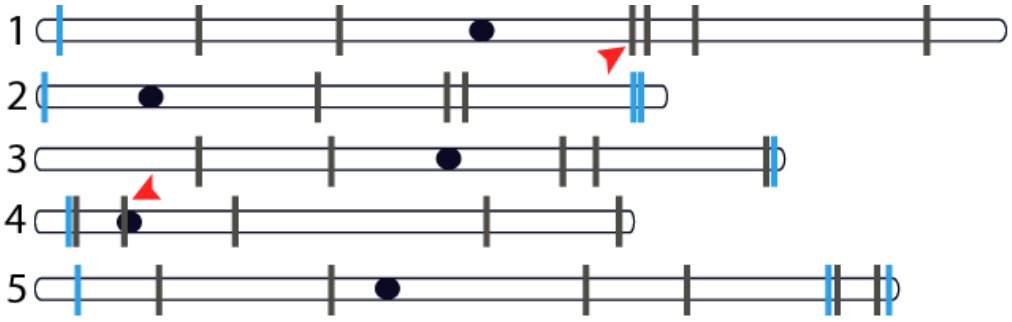


Figure 6-5: New telomeres form throughout the tetraploid Arabidopsis genome. Insertion sites for DNTF events with the 2.6kb TRA were determined for 2X and 4X wild type plants by TAIL-PCR. Insertions were mapped along the five non-homologous chromosomes (2X, blue lines; 4X, black lines). Red arrowheads denote DNTF events resulting in a 20Mb deletion and a centromere-adjacent telomere truncation. Dark ovals represent centromeres.

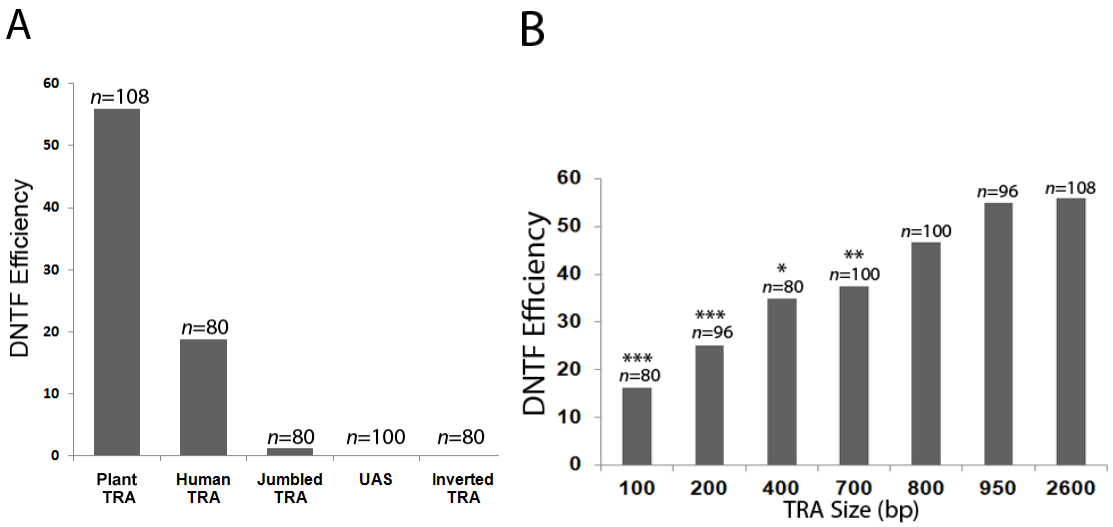


Figure 6-6. DNTF efficiency is dependent on the sequence and length of the TRA. (**A**) Wild type tetraploid plants were transformed with arrays consisting of either Arabidopsis (TTTAGGG, 950 bp), Human (TTAGGG, 750bp), jumbled (TGGTTGAT, 500bp) or UAS (CGGAGGAGAGTCTTCCG, 600bp) repeat sequences cloned into the pBGW plasmid. The frequency of DNTF was determined by PETRA using primers targeting the pBGW plasmid backbone. (**B**) TRAs of various lengths were transformed into wild type plants and DNTF events were scored by PETRA. Efficiency was calculated as described in Fig. 2D. *** indicates a p-value < 0.0001, ** < 0.001, * < 0.01.

DNTF requires a properly oriented TRA comprised of 100 bp of TTTAGGG repeats

We investigated the sequence requirements for DNTF. Orientation proved to be important, as 4X *Arabidopsis* plants transformed with a 900bp TRA in an inverted orientation failed to promote DNTF (Figure 6-6A). Altering the repeat sequence was also detrimental (Figure 6-6). Only 1% (1/80) of plants transformed with a 500bp jumbled TRA (TTGATGG)n showed DNTF. Similarly, 600bp of the UAS repeat (CGGAGGAGAGTCTTCCG) did not produce any DNTF events. In addition, DNTF was detected in ~18% (15/80) of plants transformed with a 750bp TRA consisting of the vertebrate telomere repeat (TTAGGG)n, in contrast to 37% (37/100) in plants transformed with 700 bp of the *Arabidopsis* repeat (Figure 6-6A and Figure 6-6B).

To establish an optimal TRA length for DNTF, derivatives of pWY86 with varying amounts of TTTAGGG repeats were transformed into 4X *Arabidopsis* (Figure 6-6B). Previous studies indicate that 1kb represents a critical length threshold for *Arabidopsis* telomeres (Heacock et al. 2004). Below this size, telomeres begin to be recruited into end-joining reactions. In *tert* mutants (which harbor a null mutation in the catalytic subunit of telomerase), the smallest TRA detected with an intact G-overhang is ~300bp (Heacock et al. 2004). In 4X WT, even the smallest TRA we tested, corresponding to 100bp, initiated DNTF, albeit at a substantially reduced rate relative to a 900 or 950bp TRA (16% vs 54%, respectively) (Figure 6-6B) (Table 6-1). New telomeres recovered from plants bearing a 100bp TRA ranged from 650bp to 1.9kb (avg=1.1 kb) (Table 6-2), indicating that up to 1.8kb of telomere repeats were added to

the nascent terminus in a single plant generation. A similar trend was observed with other TRAs shorter than 1kb (Table 6-1).

A small number of the plants transformed with the 100bp TRA contained new telomeres substantially shorter than 1kb (Figure 6-7A, lines 1 and 4). We suspected that such telomeres would be prone to end-joining reactions. Telomere Fusion PCR (TF-PCR) (Heacock et al. 2004) was performed using primers to test for sister chromatid fusions with the new telomere as well as fusions between the TRA and endogenous telomere 3L (Figure 6-7). TF-PCR products were generated in reactions with lines 1 and 4, but not when new telomeres were longer than 1kb (Figure 6-7B, lines 2 and 3). As expected, the endogenous 3L telomere did not engage in extensive end-joining reactions with the new TRA (Figure 6-7B, 2 + 3L). We conclude that nascent telomeres must exceed a critical 1kb size threshold to avert end-to-end chromosome fusions.

Inactivation of telomerase increases the frequency of DNTF.

We used a genetic approach to examine the role of telomerase in converting a TRA into a functional telomere. Second generation (G2) *tert* mutants were made tetraploid and then transformed in the next generation with pWY86. The frequency of DNTF was monitored in two separate transformation experiments. On average, 72% of the transformants displayed DNTF (T1=68/96 (70%), T2=53/72 (74%)), a statistically higher fraction than 4X wild type plants (56%) (p value $\leq$.01) (Figure 6-8A). Southern blot analysis confirmed that the average transgene copy number was the same (or lower) in *tert* compared to wild type transformants (Figure 6-9). Thus, the elevated frequency of

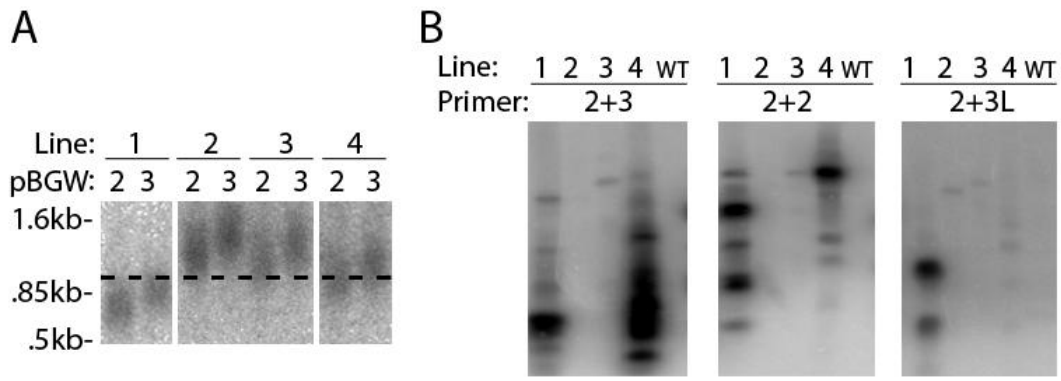


Figure 6-7. TRAs shorter than 1kb are prone to end-joining reactions.
(**A**) PETRA results are shown for four lines transformed with a 100bp TRA using PBGW-2 and -3 primers, which target a unique sequence in the T-DNA 250 or 420bp, respectively, upstream of the TRA. At least one telomere in lines #1 and #4 is shorter than 1kb (dashed line). (**B**) TF-PCR results from the lines analyzed in (**A**). WT indicates a control TF-PCR reaction performed with an untransformed 4X WT control. Primers PBGW-2 and PBGW-3 were used for the reactions on the left, and only primer PBGW-2 was used on the right. A combination of 3L and PBGW-2 primers were used to test for non-sister chromatid fusions.

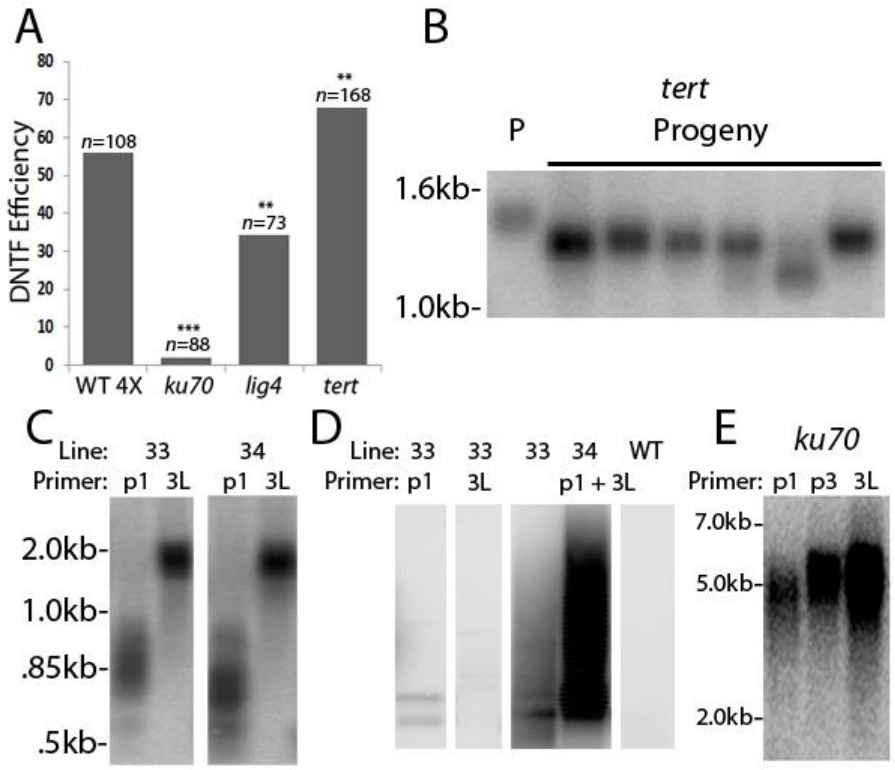


Figure 6-8. Establishment of a new telomere is dependent on Ku and Lig4, but not on telomerase.
(**A**). DNTF efficiency was monitored in the 4X mutant Arabidopsis lines shown. Efficiency was calculated as in Fig. 2D. (**B**) PETRA results for DNTF in a T1 tert deficient parent (P) and several T2 progeny. (**C**) PETRA results for two 4X tert lines (33 and 34) showing new telomeres shorter than 1kb. PETRA products for the endogenous 3L telomere are shown for comparison. (**D**) TF-PCR results for lines 33, 34 and WT. Reaction were conducted with a single primer (p1 or e3L) or with both primers. (**E**) PETRA results for a 4X ku70 transformant using either p1, p3 or e3L.

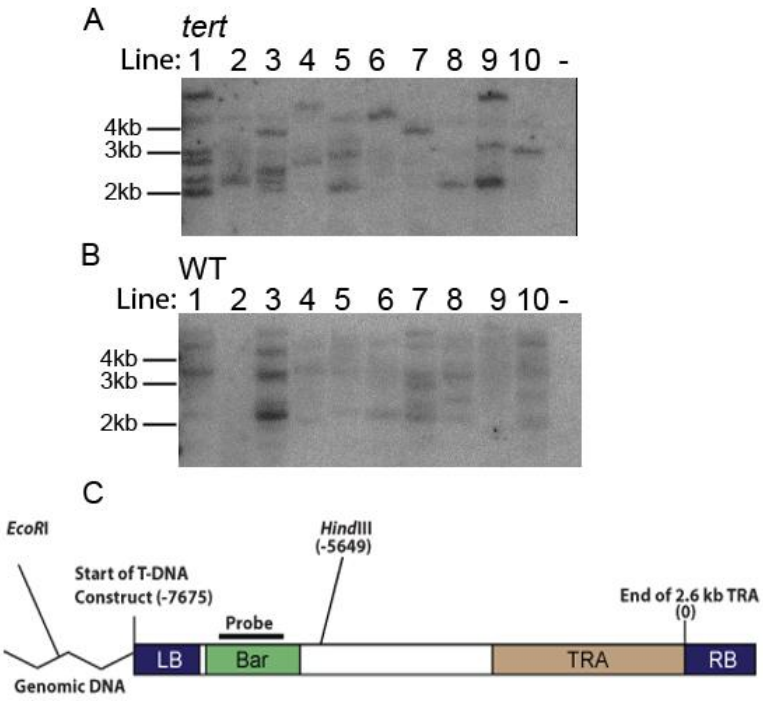


Figure 6-9. 4X *tert* acquired the same number of T-DNA insertions as 4X wild type. DNA gel blot analysis of tert (**A**) and wild type (**B**) samples digested with HindIII and EcoRI and hybridized with a probe for the BAR gene. EcoRI cuts within the adjacent genomic DNA (see map in **C**), whereas HindIII cuts downstream of the BAR gene. The negative (-) control is 4X WT without a transgene. Numbers indicate nucleotide positions within the PWY86 plasmid.

DNTF in 4X *tert* mutants does not reflect increased T-DNA integration, but rather argues that telomerase is inhibitory to DNTF.

As for endogenous telomeres, telomerase was needed to maintain the nascent telomere once it is established. In first generation (G1) *tert* mutants, endogenous telomeres are approximately 1kb shorter than wild type, and then decline by ~200-300bp each plant generation thereafter (Riha et al. 2001). Similarly, in 4X *tert* transformants (G3 for *tert*), the average TRA associated with a nascent telomere was ~1.5kb (Table 6-3), corresponding to the loss of 1.1kb from the 2.6kb pWY86 TRA. In the next generation, the new telomere decreased in size by ~200bp (Figure 6-8B).

Like 4X wild type transformants, a subset of the *de novo* telomeres formed with the 2.6kb TRA in 4X *tert* were significantly shorter than 1kb (Figure 6-8C). Only a low level of sister fusions were detected with newly formed telomeres (Figure 6-8D, p1 alone). Consistent with previous analysis of *tert* mutants, no fusions were detected with the 3L telomere control (3L alone) (Heacock et al. 2004). Unexpectedly, however, TF-PCR products were observed in reactions targeting the new telomere and 3L (p1 + 3L). We conclude that a short, unstable TRA in a telomerase-negative setting is capable of recruiting a fully capped and functional telomere into an end-joining event.

Table 6-3. Range and average length of TRAs recovered from different genetic backgrounds

Genotype	Initial TRA (bp)	TRA recovered	
		Range (bp)	Avg (bp)
Wild Type	2600	950-4100	2562 +/- 574
G4 *tert*	2600	650-2050	1529 +/- 460
G4 *tert* (endogenous)	N/A	670-2600	1466 +/- 657
lig4	2600	600-3050	1508 +/- 722

NHEJ machinery is required for DNTF in Arabidopsis

DNTF would appear to be in direct competition with DNA repair proteins, since factors necessary to authenticate the T-DNA as a telomere seed must displace or compete with components required for T-DNA integration. The situation may be more complex, given studies in yeast showing that Ku, a core component of NHEJ machinery, promotes telomerase recruitment at DSBs and at the same time protects natural telomeres from chromosome fusion (Stellwagen et al. 2003; Wang et al. 2009). To test how NHEJ components influence new telomere formation in a multicellular eukaryote, we transformed pWY86 with a 2.6kb TRA into tetraploid plants carrying a null mutation in either *KU70* or DNA ligase IV (*LIG4*).

Unexpectedly, 4X *lig4* transformants showed a statistically significant decrease in DNTF events relative to 4X wild type plants, 26/73 (36%) versus 61/108 (56%) (Figure 6-8A). The value does not reflect decreased integration of pWY86 since T-DNA integration does not require LIG4 or KU (Friesner and Britt 2003). Moreover, we assayed DNTF in 4X *lig4* transformants resistant to BASTA, and thus bearing an integrated T-DNA. The average length of a new telomere in 4X *lig4* mutants was 1.5kb, 1.1kb shorter than the 2.6kb TRA in pWY86 (Table 6-3). Since LIG4 does not make a significant contribution to telomere maintenance in *Arabidopsis* (Heacock et al. 2007), the data suggest that the TRA was subjected to nucleolytic digestion or deletional recombination prior to becoming a fully capped telomere.

An even more dramatic decrease in DNTF was observed in *ku70* 4X mutants. Only 2% (2/88) of the transformants formed new telomeres (Figure 6-8A). In

Arabidopsis, endogenous telomeres are grossly extended in the absence of Ku (Bundock et al. 2002; Riha et al. 2002). PETRA revealed that the new telomeres formed in these two lines were elongated to approximately the same extent as the endogenous 3L telomere, with the addition of ~2.5kb in one generation (Figure 6-8E). These results show that if a telomere can form in the absence of Ku, it is subjected to the same length regulation as endogenous telomeres. We conclude that NHEJ components directly or indirectly promote DNTF at chromosome breaks, and further that Ku plays an additional, specialized role in new telomere formation.

Discussion

Conversion of a DSB into a fully capped telomere by DNTF is a potentially lethal event, leading to gene loss and genome instability. Thus, cells must evolve mechanisms to strictly control DNTF. Emerging data from budding yeast reveal that the ATR ortholog, Mec1, promotes genome integrity by negatively regulating DNTF as part of the DNA damage response (Lydeard et al. 2010; Makovets and Blackburn 2009; Zhang and Durocher 2010). Factors that modulate DNTF in multicellular eukaryotes are largely unexplored, but as this study illustrates the flowering plant *Arabidopsis* is poised to fill this gap in understanding. For example, unlike vertebrate models, null mutations in ATR or telomere capping proteins are viable in *Arabidopsis* (Watson and Riha 2010). Furthermore, as demonstrated here, tetraploid *Arabidopsis* has the necessary genome buffering capacity to reveal fundamental insights into the mechanism of DNTF.

Remarkably, up to half of the 4X Arabidopsis transformants we tested acquired

truncated chromosomes capped by new telomeres. This represents a five-fold increase in DNTF events relative to diploid *Arabidopsis* and a much higher frequency than in yeast (<1%) (Kramer and Haber 1993), human embryonic fibroblasts (<2%) (Barnett et al. 1993), or maize (~9%) (Yu et al. 2006). The incidence of DNTF events in 4X *Arabidopsis* is comparable to mammalian cancer cell lines (40-60%), which are characterized by abrogated cell cycle checkpoints and rampant aneuploidy (Nigg 2001). In normal diploids, including *Arabidopsis*, DNTF events are almost uniformly recovered near chromosome ends (Barnett et al. 1993; Yu et al. 2006). In contrast, DNTF events arise throughout the entire 4X *Arabidopsis* genome, an outcome that does not reflect an increased T-DNA integration in tetraploid plants. Rather, the genetic redundancy of the tetraploid genome appears to provide a less stringent filter for telomere-mediated chromosome truncation.

A critical length threshold for new telomeres

Arabidopsis can establish a telomere with a TRA of only 100 bp. Indeed, DNTF events may occur with little to no TRA seed sequence. *Arabidopsis* telomerase extends primers lacking any complementarity to the telomerase RNA template *in vitro* by aligning the 3' terminus at a "default" position within the RNA (Fitzgerald et al. 2001). Moreover, in yeast and mammals telomeres can be formed in the complete absence of a TRA (Flint et al. 1994) via microhomology between the nucleotides at the 3' terminus of the break site and the telomerase RNA template (Stellwagen et al. 2003). DTNF at the α-globin locus in humans and in mouse embryonic stem cells is thought to proceed by such a mechanism (Sprung et al. 1999; Wong et al. 1997).

By exploiting a PCR strategy devised to follow the fate of individual *Arabidopsis* telomeres, we discovered that the TRA must be elongated above a critical 1kb length threshold to establish a fully capped telomere that prohibits end-joining reactions. The average length of new telomeres formed with a 100bp TRA was 1.1kb. Moreover, telomere fusion events were detected with TRAs shorter than 1kb, while longer TRAs were immune. The molecular switch underlying this length threshold is unknown, but it may represent the minimal TRA required to recruit a sufficient number of telomere capping proteins or to assemble into a protective secondary structure such as a t-loop.

Competition for the nascent chromosome terminus

Once a TRA is exposed, we speculate that it is engaged by multiple competing pathways (Figure 6-10). As in mammalian cells (Okabe et al. 2000), we found that telomerase is not required to establish a new telomere when the break occurs adjacent to a TRA (Figure 6-10, steps 1-3). However, telomerase is needed to maintain the integrity of the new telomere in successive generations (Figure 6-10, step 4). Strikingly, successful de novo telomeres formed from short TRAs are extended past 1kb in the first generation, and then brought into the wild type range in the second generation after transformation. Intriguingly, DNTF occurred at an even higher frequency in telomerase-deficient 4X plants than in wild type 4X. One explanation is that telomerase interferes with DNTF by competing with proteins needed to form a protective cap and thus stabilize the nascent terminus (Figure 6-10, steps 2 and 3a).

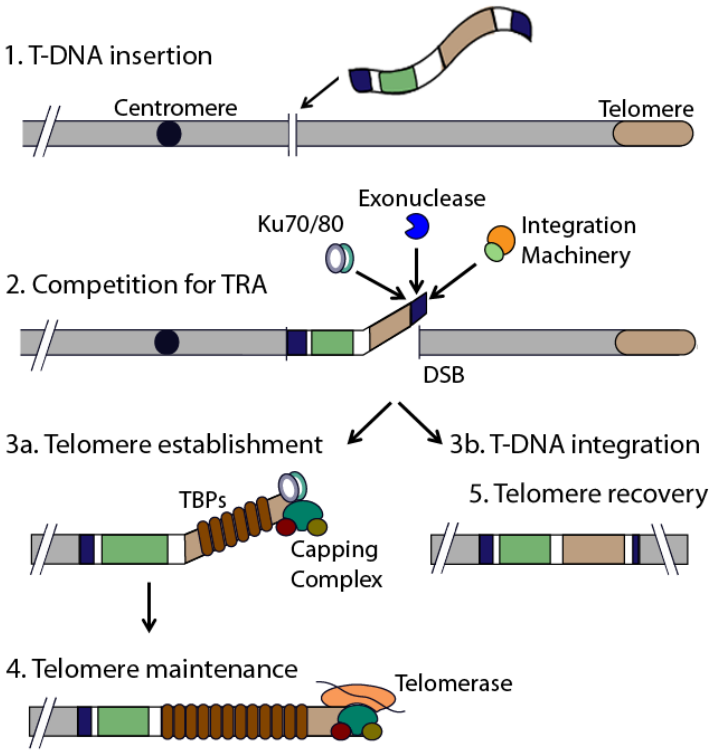


Figure 6-10. Model for DNTF establishment and maintenance in Arabidopsis. After the first step of T-DNA insertion (Step 1), there is a competition for the TRA by Ku, T-DNA integration machinery, and exonucleases at the DSB formed at the 3' end of the T-DNA (Step 2). T-DNA integration will occur if integration machinery out-competes Ku (Step 3b). If Ku out-competes the integration machinery, controlled exonucleolytic processing exposes the TRA (Step 2). A new telomere is established by association of double-strand telomere binding proteins (TBPs) and the CST capping complex (Step 3a). After telomere establishment, the new telomere must be maintained by telomerase (Step 4) above the critical 1kb length threshold to prevent end-joining reactions.

We also exploited the genetic tractability of *Arabidopsis* to investigate how NHEJ machinery affects DNTF in a multicellular organism. Unexpectedly, we discovered that both Ku and LIG4 directly or indirectly promote DNTF. NHEJ occurs by multiple routes in *Arabidopsis*, including a Ku and LIG4-independent pathway. Neither of these factors is required for T-DNA integration (Gallego et al. 2003; Heacock et al. 2007; Li et al. 2005; van Attikum et al. 2003). The frequency of DNTF events in 4x *lig4* mutants was significantly reduced and the new telomere tracts were shorter than in 4X wild type plants. Notably in plants doubly deficient in LIG4 and TERT, extreme nucleolytic degradation is observed prior to end-joining (Heacock et al. 2007). Therefore LIG4 may temporarily sequester the TRA from nucleolytic attack, providing additional time for telomere capping proteins to engage the terminus (Figure 6-10, step 2).

Finally we discovered that Ku plays a critical role in DNTF in *Arabidopsis*. Unlike LIG4, Ku is an essential component of the chromosome terminus, functioning in both telomere length regulation and protection of the telomeric C-strand (Bertuch and Lundblad 2003; Wang et al. 2009). Bertuch and colleagues propose a two-faced model for yeast Ku that explains its dual functions in NHEJ and telomere maintenance. In this model, Ku differentiates between telomeric and non-telomeric DNA based on the orientation of its two molecular faces (Ribes-Zamora et al. 2007). Accordingly, Ku interaction with the TRA on the T-DNA could define this region as a telomere instead of a double strand break. In this view, full integration of the T-DNA into the chromosome would be favored over DNTF (Figure 6-10, steps 3a and 3b). In contrast to yeast, Ku does not associate with the TER1 telomerase RNP in *Arabidopsis* (Cifuentes-Rojas et al.

2011), and instead acts as a potent negative regulator of telomere length (Riha et al. 2002). Thus, Ku could potentially play an indirect role in promoting DNTF. For example, in the absence of Ku, the ultra-long endogenous telomeres may sequester double-strand telomere binding proteins, preventing the transgenic TRA from being established as a functional telomere.

DNTF and chromosome engineering

Our findings indicate that natural polyploids can be manipulated as a platform for chromosome engineering and plant breeding through telomere truncation. Given the high frequency of DNTF and a semi-high throughput method for identifying these events, it should now be possible to recover chromosome truncations at a desired location. With different selectable markers, a streamlined chromosome could be created by multiple truncation events and then reintroduced into a diploid by conventional genetic crosses (Yu et al. 2007) or through centromere-mediated genome elimination (Ravi and Chan 2010). Finally, since truncated chromosomes can be transmitted to progeny, the consequences of such events could be examined over several generations.

CHAPTER VII

CONCLUSIONS AND FUTURE DIRECTIONS

Telomeres are an ancient evolutionary mechanism for masking the ends of chromosomes and preventing them from being recognized as double-strand breaks (DSBs). Gradual erosion of telomeric DNA over time is of minimal significance in light of the immediate and devastating consequences that arise from a DSB. Therefore, the amount of regulation and chicanery that goes into masking telomeres from being recognized as a DSB is remarkable.

Support for the importance for proper telomere capping over telomerase abounds, as most cells in a multicellular organism undergo daily metabolism without telomere extension by telomerase. Some organisms, such as Arabidopsis, can exist for multiple generations in the absence of telomerase, yet suffer severe genome instability in the absence of telomere capping components (Riha et al, 2001; Song et al, 2008; Surovtseva et al, 2009). Thus, the proper sequestration of the chromosome end into a protective nucleoprotein complex that will hide it from unfriendly enzymatic and sensing activities is crucial.

How do telomeres do this? The simple answer would be to say that telomeres are isolated complexes with a unique protein composition that is dedicated soley to the chromosome end. However, telomere-associated associated proteins are not always associated with the telomere, underscoring how complicated life really is at the periphery of genomes.

As I have alluded to throughout this dissertation, telomeres associate with several components of the DDR machinery (ATM and ATR, Ku, and MRN/X), and these factors are necessary for telomere maintenance. In return, telomere components and DDR machinery work to prevent DSBs from being converted into telomeres. It is an uneasy alliance, but one that evolution is finding new ways to strengthen.

In this dissertation, I highlighted some of the novel mechanisms used by *Arabidopsis thaliana* to regulate telomerase during DNA damage and at DSBs (Chapters II, VI). In addition, I presented a novel assay in which Arabidopsis can be used to test for genetic elements that regulate inappropriate telomere addition at DSBs (Chapters VI). The utilization of TER2 as an inhibitor of telomerase activity may be a recent evolutionary innovation, and was discussed in Chapter II and III. Finally, the protein components at the chromosome end are also rapidly evolving in Arabidopsis. I discussed the initial characterization of AtPOT1b and POT1c, and their differing roles on and off the chromosome end (Chapter IV and V). In this Chapter, I will discuss the conclusions and future directions that have arisen from this research.

A telomerase RNA, TER2, regulates telomerase in response to DNA damage

During the course of identifying and characterizing the telomerase RNA subunit in Arabidopsis, our lab found three TER isoforms, TER1, TER2, and a shorter form of TER2, TER2s. Intriguingly, over-expression of TER2 *in vivo* leads to a strong inhibition of telomerase activity and decrease in telomere length. A mutation within the template of the over-expressed TER2 molecule allowed us to determine that this molecule was part

of a telomerase complex that did not efficiently incorporate telomere repeats onto the chromosome end *in vivo*. A null mutant was available for TER2 (*ter2-1*), and these mutants display an increase in telomerase activity, but no apparent change to telomere length.

Thus, *Arabidopsis thaliana* contains a regulatory telomerase RNP. When and where does the alternative telomerase RNP? By monitoring telomerase activity and TER RNA levels in seedlings exposed to the genotoxin zeocin, I found that DNA damage results in repression of telomerase. This repression is dependent on TER2, as *ter2-1* lines do not show a telomerase repression. In fact, *ter2-1* lines display a slight but significant increase in telomerase activity in the early time points of treatment. This finding shows that TER2 is required for a specific physiological response, but it does not convey a mechanism. Furthermore, this may not be the only instance of repression by TER2.

The most efficient means of prohibiting telomerase action at DSBs may not be mediated through a TERT-bound RNP. In Chapter II, we presented evidence that Ku, ATR, and AtPOT1b associate with TER2, and to varying degrees, TER2s, *in vivo*. Each of these components could be interacting separately or in concert with one another to inhibit telomerase at the site of DNA damage. I will discuss potential roles for each of them below.

A possible Ku-TER2 telomerase inhibitory complex at DSBs

Firstly, Ku interacts with the telomerase RNA subunit in budding yeast through a conserved RNA binding domain found within the Ku80 subunit (Stellwagen et al, 2003).

Strikingly, this 15 amino acid region is highly conserved among Ku80 from multiple organisms, including Arabidopsis (Figure 7-1A). This domain is not conserved in *S. pombe* Ku80, which does not interact with TER (Subramanian et al, 2008). Ku wraps around dsDNA, allowing for DNA and RNA binding to occur simultaneously (Stellwagen et al, 2003). Thus, Ku-bound TER2 could act as a binding platform at DSBs for the recruitment of other factors necessary for the stimulation or repression of DNA repair activities.

A variety of factors important for DNA repair and inhibition of telomerase could associate at Ku-TER2 foci. ATR, the master kinase, associates with TER2, perhaps directly, or indirectly with AtPOT1b (Vespa L, Jasti M, and Shippen D, unpublished data). ATR is important for the regulation of DNA repair, but also regulates telomerase activity in many eukaryotes, including Arabidopsis (Boltz et al, 2012; Yamazaki et al, 2012; Chawla et al, 2011; Moser et al, 2009). In addition, in Chapter IV, I presented data to suggest that AtPOT1b is a negative regulator of telomerase. Thus, TER2 is already acting as a scaffold for the appropriate factors to inhibit telomerase at DSBs (Figure 7-1B).

While purely speculative, this model is not without precedent. A recent explosion of new information sheds light on the role of non-coding RNAs (ncRNAs) in gene regulation and DDR. There are thousands of long, intergenic, noncoding RNAs (lincRNAs, ie, TER2) encoded within the human genome (Clark et al, 2012). Many of these RNAs have a long half-life and move between different cellular compartments. In plants, lincRNAs have been implicated in environmental stress responses and regulation

of development (Ben Amor et al, 2009; Charon et al, 2010). Thus, these "Riboregulators" provide a novel means of regulating cellular functions.

Polycomb Repressor Complexes (PRCs) load onto long non-coding RNAs (ncRNAs) that acts in trans to recruit chromatin remodelers to genes in need of silencing (Heo and Sung, 2011; Swiezewski et al, 2009). Even more recently, DSBs were found to stimulate the production of small RNAs from the region immediately around the break site (Wei et al, 2012). These DSB-induced RNAs (diRNAs) were critical for efficient DSB repair. Ultimately, diRNA production was dependent on the kinase ATR. Perhaps a connection lies between ATR, TER2, and diRNAs? More experimental data is needed to address the role of Ku and TER2 in DNA damage.

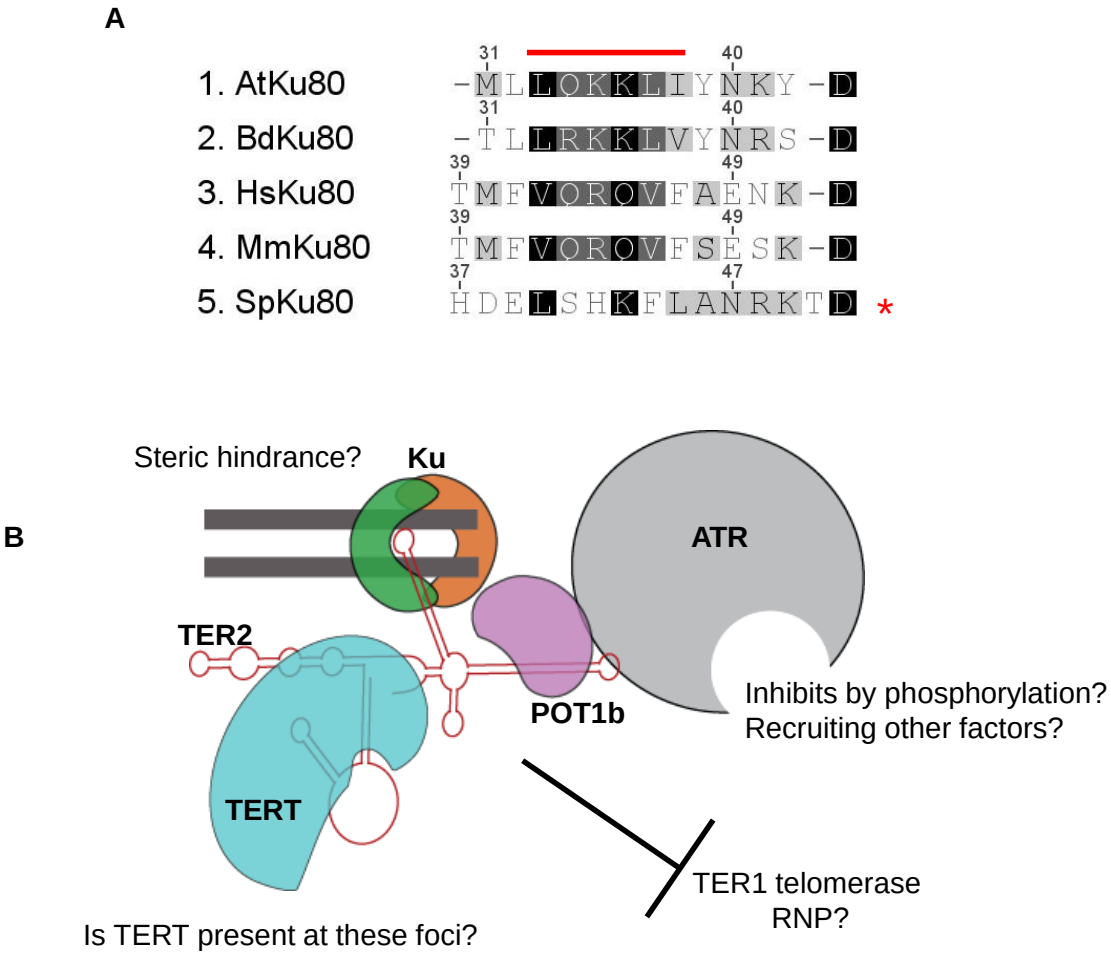


Fig 7-1. TER2 as a scaffold for inhibitory factors of telomerase. (**A**) Amino acid alignment of the predicted RNA binding motif of Ku80. At= *Arabidopsis thaliana*, Bd= *Brachypodium distachyon*, Hs=*Homo sapiens*, Mm=*Mus musculus*, Sp= *Schizosaccharomyces pombe*. The minimal RNA binding domain is indicated by the red bar above the alignment. The asterisk denotes the Ku80 protein that does not bind TER (*S. pombe*). (**B**) A model for inhibition at DSBs by the TER2 RNP complex. Shown are protein constituents that interact with TER2 *in vivo*. In this model, Ku recruits TER2 to DSBs through Ku80. TER2 then serves as a scaffold for the localization of POT1b , ATR, and potentially TERT to DSBs. This is likely a very dynamic process, and thus a complex in this form might interact only transiently at DSBs to block telomerase access until NHEJ components can be assembled.

To address the role of a Ku/TER2 complex in the DNA damage response, I propose the biochemical purification of TER2 RNP complexes from plants treated with DNA damaging agents. The increased number of TER2 molecules seen during DNA damage may assemble into a complete TER2 RNP, or with only a subset of components. If a Ku/TER2 RNP forms in response to DNA damage, then immunoprecipitation of Ku should reveal increased association with TER2 or TER2s. By subjecting the IP to western blotting, one could determine if TERT and POT1b also associates with Ku in response to DNA damage. Notably, Ku80 recognizes a fairly simple 48nt stem-loop structure that conceivably could be contained within other ncRNAs. Thus, deep-sequencing RNAs that co-purify with Ku after DNA damage may reveal additional regulatory RNAs.

Evidence for alternative mechanisms for inhibition of telomerase following DNA damage

Treatment of *ter2-1* lines with genotoxin revealed an increase in telomerase activity during the earlier time points, followed by a gradual decline in activity to levels seen in wild type treated seedlings. One possible explanation for this is that the population of cells where telomerase is presumably most active, the stem cell niche, has been obliterated. An alternative interpretation is that newly synthesized telomerase components are sequestered from one another into different compartments of the cell during DNA damage.

Evidence supporting this latter model comes from human cells, where oxidative stress and DNA damage causes hTERT to be excluded from the nucleus and accumulate in the cytoplasm. hTERT is also translocated to the mitochondria where, in an unknown manner, it makes the cell more resistant to genotoxic stress (Passos et al, 2007; Saretzki G, 2009; Ahmed et al, 2008). These same reports also demonstrated that TER was not associated with this mitochondrially bound TERT, reinforcing the notion that these two entities are prevented from forming RNP complexes during DNA damage.

Is TERT nuclear exclusion conserved in Arabidopsis? To test this, nuclear and cytoplasmic protein extracts coule be prepared and monitored for TERT localization after DNA damage. If preliminary experiments were promising, then these fractions could be tested for the presence of TERT-bound TER molecules. In addition, factors assisting in exporting TERT out of the nucleus could be identified from these same IPs.

A Ku/TER2s/POT1b complex at telomeres

Above, I have focused primarily on the role of Ku, TER2, and ATR at DSBs, but DSBs are limited in number and most likely do not reflect the bulk of TER2's roles. Ku regulates C-strand resection and regulates telomere length in Arabidopsis. Contrary to its function in human or budding yeast, Ku actually restricts telomere addition. In the absence of Ku, telomeres are rapidly over-extended in a telomerase dependent manner. However, this does not reflect an increase in telomerase activity. Instead, this likely indicates an inhibitory complex is removed from telomeres in a *ku* background.

Based on preliminary data presented in Chapter IV, I propose that this Ku-dependent inhibitory complex likely consists of POT1b and TER2s. There are some caveats to the gross telomere uncapping phenotype in the *ter2pot1b* double mutant (Chapter IV). However, the physiological and telomere phenotypes observed in this background are almost certainly caused by POT1b and TER2.

If the dramatic phenotype observed in the *pot1b-1/ter2-1* background is indeed derived from mutations in these two genes, the data suggests that chromosome-end protection is in part mediated by a telomerase RNP complex (Figure 7-2). This proposal is not without precedent. Independent *TERT* or *TER* mutants in *Candida albicans* accumulate long G-overhangs, consistent with extensive nucleolytic degradation (Hsu et al, 2007). Furthermore, in budding yeast, telomerase is present in a Ku-dependent manner in the G1 phase of the cell cycle-when telomeres are not being extended. Thus, telomerase association with the telomere may be important for functions outside of DNA replication.

A highly speculative model for a Ku/POT1b/TER2s capping complex is that Ku localizes to long telomeres and prevents further telomere extension. Perhaps Ku/Ter2/POT1b keep G-overhangs short at these telomeres, inhibiting telomerase by removing its substrate. Anecdotal evidence comes from the Riha lab, which reports that Ku binds to and maintains "blunt-ended" telomeres. A switch would occur at short telomeres, perhaps due to telomere architecture or a "counting mechanism," involving double strand telomere-binding proteins analagous to the model proposed in budding yeast (Shore and Bianchi, 2009). In this case, CST would come to short telomeres and recruit an active telomerase through CTC1-POT1a interactions, leading to telomere extension (Figure 7-2).

Alternatively, Ku/POT1b/TER2s may bind to telomeres during G1 and G2 phases of the cell cycle, while CST would be bound during the S phase. A prediction of this model is that lines deficient for both Ku and CTC1 would show synergistic defects associated with chromosome-end deprotection and telomere loss. Preliminary data suggests that this is indeed the case (Song X, Boltz K, Leehy K, Shippen D, in preparation).

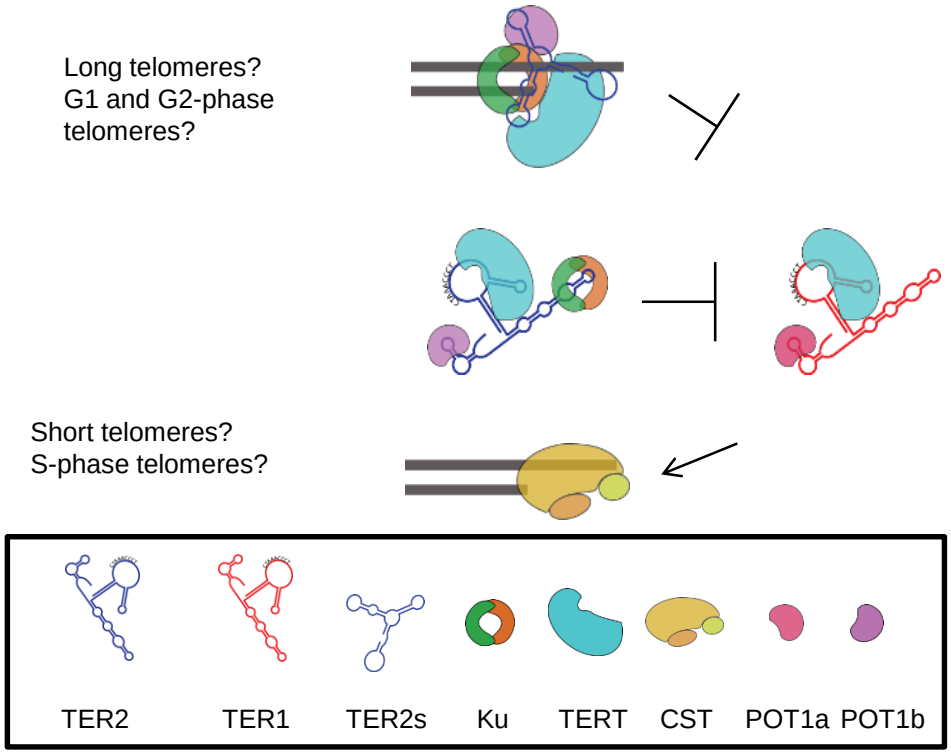


Fig 7-2. A proposed model for a Ku/POT1b/TER2 telomere capping complex. TER2 acts as a natural competitive inhibitor of the TER1 telomerase RNP. TER2s through Ku and POT1b, may represent an evolutionarily conserved telomere capping complex, perhaps at long telomeres where telomerase extension is not required. Alternatively, this TER2/POT1b/Ku/POT1b/TER2/X (where X denotes unknown components, thus KBTx), could be bound at telomeres during the G1 and G2 phases of the cell cycle. A switch to the CST complex would be necessary during the S phase in order to extend and replicate telomeres.

To fully test this model, several experiments are needed. Immunolocalization of Ku and a tagged TER2 in Arabidopsis nuclei are necessary to determine if these molecules co-localize to telomeres. This experiment could be performed in different genetic backgrounds, such as *tert* and *pot1b*, to determine the components necessary to localize a Ku/POT1b/TER2s complex to telomeres. Genetic experiments involving a *stn1pot1b* double or *stn1pot1bter2* triple would address whether these two complexes protect telomeres independently of one another.

If it turns out that the causal agent in the pot1b/ter2 mutants is neither of these two genes, map-based cloning can be used to to map the causative locus. In addition, a polymorphism in either the Col-0 or Ler-0 may be contributing to this phenotype in conjunction with TER2 or POT1b. To address this genetically, crosses should also be made between *ter2-1* (Col-0) and a wild type Ler-0 line. Reciprocal crosses involving *pot1b-1* (Ler-0) and wild type (Col-0) have already been produced, with no obvious telomere phenotype. The *ter2-1*/wild type (Ler-0) crosses are underway, and should help clarify the role of POT1b and TER2 in chromosome-end protection.

Probing the TER2 IS function and self-splicing activities: lessons from Chapter III

In Chapter III, I addressed the evolution of the TER2 IS in *A. thaliana*, and similarities between TEs and TER2 IS. Binding studies show that TERT prefers binding to TER2 10-fold over TER1 (Chapter II). These, and other data, suggest that full length TER2 is inhibitory due to its ability to sequester telomerase into complexes that cannot

interact productively with the chromosome end. We now have a host of information, from the 1,001 Arabidopsis genomes project, to test this model of inhibition.

Of the three ecotypes naturally missing the IS (Baa, No, and Ler), No and Ler are currently growing in the lab. The function being performed by TER2s is still present in these ecotypes, because they express TER2s at levels similar to Col. What is the response of these ecotypes to DNA damage? Additionally, would exogenous Col-0 TER2 splice in the Ler-0 or No-0 backgrounds, and would it inhibit telomerase activity? We have no clear understanding of the degree to which self-splicing occurs *in vivo*, thus placing full length TER2 back into ecotypes with just TER2s could help answer this question.

Sixty other ecotypes showed high degrees of variation within the IS, providing the lab with candidates to probe for *in vitro* and *in vivo* processing. Once a consistent *in vitro* splicing assay has been developed, these RNAs can be screened to delineate the regions of the RNA necessary for splicing. Concurrently, RNA can be extracted from these ecotypes and tested for the presence of TER2s. It would be interesting to find mutants that could not self-splice *in vitro*, but were still processed *in vivo*.

Transposable elements and their relationship to the TER2 IS

The finding that the TER2 IS is similar to transposable elements (particularly the MITEs) provides an interesting glimpse into the evolution of TER in *A. thaliana*. The presence of the IS provides a model for how the TER genes were duplicated, and why they are not present in syntenic regions within *A. lyrata*. Transposable elements, by their

very nature, are mobile. During transposition, they can “transduplicate” DNA around them, in essence moving the DNA around them to a new location. In addition, the small Class II DNA transposons prefer to embed themselves adjacent to active genes.

This seems to have happened to TER1 and TER2. I propose that somewhere on the lineage leading to *A. thaliana*, an IS inserted within a single copy ancestral TER molecule and then transposed to new locations within the genome (Figure 7-3A). This created two copies of a TER with an IS. During the course of evolution, the IS was removed from TER1. This event also occurred at the TER2 locus from three different ecotypes, leaving a TER molecule remarkably similar to TER1.

TER may still be associated with an IS in *A. lyrata*, accounting for our difficulty in identifying this locus. However, if the IS is not linked to AlTER, then the IS may have confounded efforts to find AlTER. Our searches have been for an ~700nt TER when instead the RNA may be much smaller.

If the above model is true, and TER duplication and migration through the genome was facilitated by IS insertion, then the genomic landing spot chosen by the IS may have resulted in the incorporation of non-TER sequence into TER1. Recent data indicates that TER1 landed in the 5' UTR of Rad52 of *A. thaliana* (Samach et al, 2011). Transcription of TER1 occurs in the same direction as Rad52 and terminates just prior to the beginning of the 3rd exon (Figure 7-3B). TER2 landed in the 5' UTR of an uncharacterized gene, but it is oriented in the opposite direction. The IS is removed from TER2 to produce TER2s, an RNA scaffold that is still capable of interacting with TERT, POT1b, and Ku. Therefore, I predict that the non-core regions of TER1 are accidental incorporations. Thus, TERs from other members of the Brassicaceae family may be much smaller in size and still retain function.

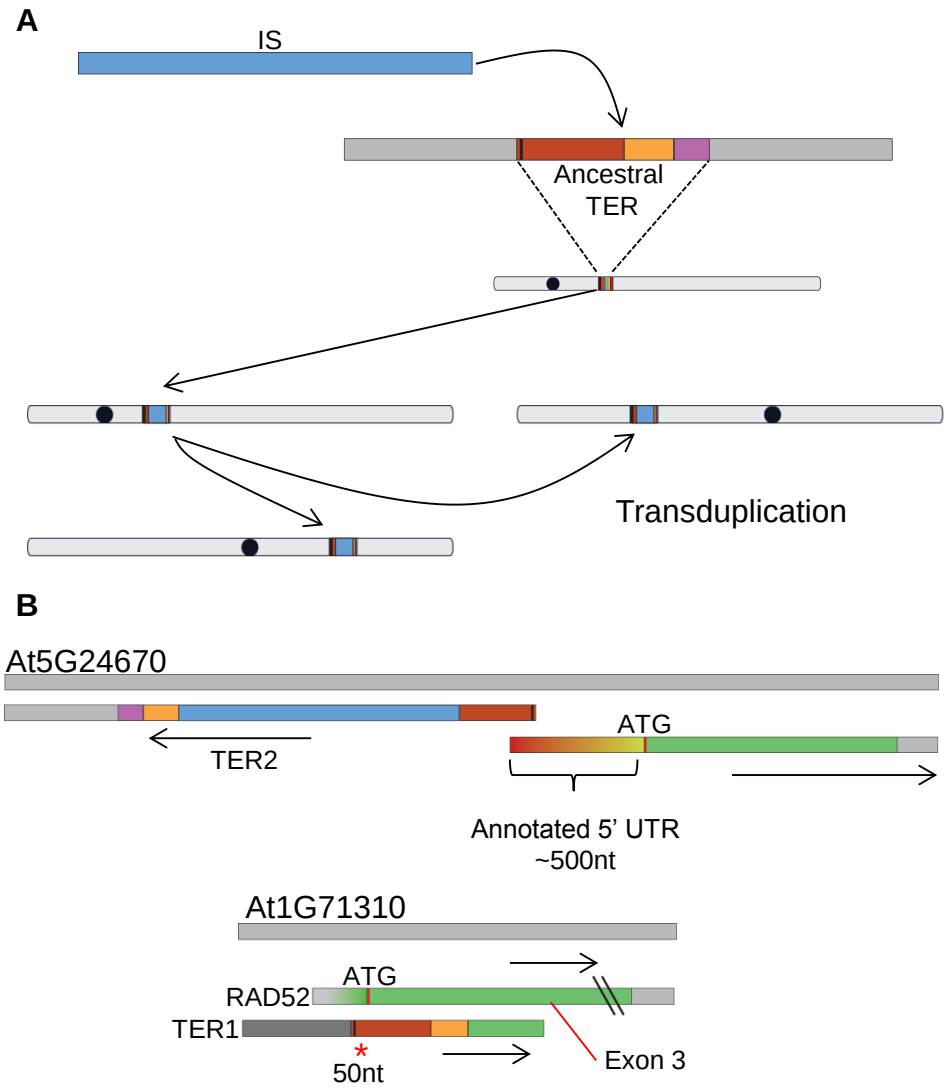


Fig 7-3. The Arabidopsis TERs may have arisen from a transduplication event. (A) Model for IS insertion and transduplication of the TER locus. IS inserted into the ancestral TER. Then, transduplication of the IS and the DNA around it (TER) resulted in TERs being duplicated and relocated to different chromosomes in *A. thaliana*. (**B**) Schematic diagram of the TER2 (top) and TER1 (bottom) loci. TER2 is oriented in the opposite direction of the gene it is embedded near. The template region of TER2 lies within the extreme 5' UTR of the adjacent gene, ~500nt from the ATG. TER1 is oriented in the same direction as the gene it is embedded within, RAD52. The template region is 50nt from the ATG of RAD52 (red asterisk). TER1 terminates near the intron 2/ exon 3 junction of RAD52.

POT1b and its role in RNA maturation

Another aspect of TER2 maturation is the removal of the 3' 36nts. The 3' end is part of a ~100nt intergenic region that is completely conserved among the different *A. thaliana* ecotypes. In addition, this region is also found at least twice in the *A. thaliana* genome, and is conserved in *A. lyrata*. It is not associated with the ISL elements, suggesting an incidental association with TER2. *In vitro* data suggests that this region is required for self-splicing of the TER2 IS (Hernandez A, Shippen D, unpublished data). Is this region, like the IS, self-spliced, or is it dependent on cellular RNA maturation machinery. Two interesting points should be made here, one involving Dyskerin, and the other, POT1b.

Dyskerin is important for RNA maturation and associates with both TER1 and TER2 (Appendix, A-1; Chapter II). However, Dyskerin does not interact with TER2s, suggesting the binding site has been removed. Dyskerin is part of a class of proteins that bind an RNA structural motif called an H/ACA box. TER molecules from other species retain an H/ACA box near the 3' end. A putative H/ACA box has been identified *in silico* at the 3' end of TER2. Importantly, this region lies within the 3' region that is removed upon TER2s maturation. This would explain lack of Dyskerin binding to this factor, but still needs to be experimentally confirmed.

Dyskerin performs RNP maturation in cajal bodies, subcellular compartments associated with the nucleolus. Thus, cleavage of the 3' could be important for the release of a TER2s molecule into the nucleus. Mapping the nulear localization of TER2s

molecules with and without the 3' end should be performed to determine if the 3' region in critical for nucleolar retention.

POT1b associates with both TER2 and TER2s *in vivo* and *in vitro*. However, immunoprecipitation of POT1b showed an enrichment of TER2s (Chapter II). This finding indicates that POT1b may associates with TER2 during the TER2 maturation process. Notably, *pot1b-1* lines do not show evidence of telomere dysfunction and instead have increased telomerase activity (discussed below). In *pot1b-1* lines, TER2s retains the 3' end normally removed during TER2s maturation (Chapter IV). Thus, POT1b is necessary for proper TER2 processing. Finally, since the TER2 in this background (Ler-0) lacks the IS, we cannot make any conclusions about the role of POT1b in the TER2 splicing reaction. POT1b is likely recruiting a factor to TER2s, perhaps by assisting in RNA folding to make it accessible to splicing factors.

Based on the requirement for POT1b in TER2 3' processing, an interesting experiment would be to use POT1b for *in vivo* pull-downs. Performing this pull-down in the presence and absence of RNase would allow mass-spec mediated indentification of potential protein-protein and protein-RNA interactions. In addition, performing the POT1b pull-down in genetic backgrounds with and without full length TER2 (Col-0 and Ler-0) may identify unique factors necessary for processing of the IS. Two interesting candidates are STEP1 and Whirly, two putative RNA binding proteins which seem to regulate Arabidopsis telomerase (Kwon and Chung, 2004; Yoo et al, 2007).

The newest POT1 paralog, POT1c

The *A. thaliana* genome contains three POT1 genes. Two of these, POT1a and POT1b, are highly conserved within the Brassicaceae family (Shakirov et al, 2010). POT1c is unique to *A. thaliana*, but is highly conserved in ~460/513 currently sequenced *A. thaliana* ecotypes. The other ecotypes show some rearrangements, but still retain POT1c. In Chapter V, I presented evidence to suggest that POT1c is a functional gene in Arabidopsis, critical for telomere protection. The architecture of POT1c could have something to do with its retention. As a single OB-fold protein, POT1c is quite versatile, capable of multimerizing to broaden its function. Notably, the functions I elucidated for POT1c overlap the roles of POT1 proteins in vertebrates and yeast.

Although the Brassicaceae POT1a and POT1b hae evolved to bind RNA, they must have ss telomeric DNA binding activities. In the moss *Physcomitrella patens*, thought to represent one of the most basal plant species, the POT1 gene is single copy and encodes for a protein that binds ss telomeric DNA and protects chromosome ends from degradation and fusion (Appendix, A-2). POT1 proteins from different plant lineages show a high degree of dissimilarity. Thus, it would be interesting to test representatives from each family biochemically to determine how the function of telomere capping and telomerase regulation is evolving in the plant kingdom.

In Brassicaceae, the POT1a and POT1b duplication originates at the base of the family. Data suggests that AlPOT1a can complement loss of AtPOT1a. This is interesting, since POT1c is *A. thaliana* specific. AtPOT1b also shows a high degree of conservation among the Brassicaceae, suggesting conservation of function between *A.*

lyrata and *A. thaliana* POT1bs. Further biochemical experiments are needed to determine if Brassicaceae POT1b members perform similar functions to one another, such as TER and Ku binding.

So how does POT1c fit into this picture? If POT1c is a sub-functionalization event, which progenitor POT1 protein is POT1c using as a functional parent? The POT1c locus is clearly derived from AtPOT1a, yet AtPOT1a does not have any apparent function in chromosome-end protection. Perhaps AlPOT1a assumes two roles in TER binding and chromosome-end protection. In *A. thaliana*, the capping function is now attributed to POT1c.

As a telomere specific protein, POT1c may have acquired a function that is not shared by its immediate progenitors. This raises an important point. In the POT1c study, we were focusing on roles for POT1c in telomere biology, since POT1c mutants show such a severe phenotype. However, as a novel, single OB-fold protein, POT1c may have acquired roles outside of telomere biology. Determining POT1c interactors, either by yeast-2-hybrid, or by IP and mass spec of over-expressed protein would be a good starting point in determining if POT1c has alternative roles in Arabidopsis.

How POT1c protects the chromosome end, interactions with CST

POT1c inherited many important interaction partners from its progenitor, AtPOT1a. Preliminary data indicate that POT1c can form homo-oligomers *in vitro*. It can also interact with CTC1, and weakly with TEN1. The interaction between POT1c

and TEN1 may be due issues not relevant to this dissertation, so I will focus on possible ramifications of the interaction between POT1c and CTC1.

Knockdown of POT1c resulted in hallmark symptoms of telomere instability similar to, but not as severe as lines bearing mutations within CTC1 (Chapter V). Given the physical interaction between CTC1 and POT1c, I propose POT1c acts to stabilize the CST complex at telomeres and to regulate access of telomerase to this complex. This POT1c-CTC1 mediated telomerase regulation likely happens through a very straightforward mechanism.

Due to their similarity, there is always the potential for competition for binding partners between POT1a and POT1c. Over-expression of POT1c leads to telomeres heterogeneous in length that form very distinct subpopulations visible by PETRA analysis, similar to, but longer than *pot1a*. In addition, telomerase activity is decreased in POT1c over-expression lines, even though TER levels are unchanged (Nelson A, unpublished data). This and other data suggests that over-expressed POT1c competes with POT1a for binding sites on TER1 or CTC1, two interactions that they share.

Preliminary data indicates that POT1a and POT1c bind to the same structural motif on TER1. POT1c and POT1a also bind the same region of CTC1, again potentially placing them in direct competition with one another *in vivo*. Notably, over-expression of POT1c in a *pot1a* background does not complement the progressive telomere shortening phenotype (Nelson A, unpublished data). This observation argues that POT1c is a competitive inhibitor of POT1a, one that regulates when and where POT1a recruits telomerase to the telomere.

Experiments are in process to examine the competitive nature of POT1c. First, YFP-POT1c IPs are being tested for telomerase activity and the identity of the associated TERs. These IPs are being performed in two different backgrounds, wild type and *pot1a*. If POT1a competes with POT1c for TER binding, then POT1c IPs from *pot1a* mutants should show an enrichment of TER1. If there is no enrichment of TER1, then perhaps POT1c is acting through another pathway, one that is inhibitory to the CST complex.

This alternative complex may be the Ku/TER2/POT1b complex mentioned above. POT1c also interacts strongly with TER2 and TER2s. Additionally, POT1c RNAi performed in *pot1b-1* lines shows a dramatic and immediate drop in telomere length to well below 1kb (Nelson A, unpublished data). Again, similar to knockdown of POT1c in a wild type background, these lines did not display rampant fusions, suggesting that loss of POT1c does not trigger a conventional DNA damage response.

DNTF in Arabidopsis, preliminary data for ATR and future experiments

In Chapter VI, I established Arabidopsis as a platform for determining factors that regulate *de novo* telomere formation. In this approach, telomere repeat arrays (TRAs) of varying lengths are transformed into tetraploid Arabidopsis and DNTF frequency is monitored by a specific PETRA assay. The manuscript describing this assay, which scored a six on the Faculty 1000 list, allowed us to determine that Ku was critical for DNTF, and surprisingly, TERT was inhibitory. However, the true importance of the experimental design is in the questions that can now be answered using this approach.

Inhibition of DNTF is still very much a black box in multicellular eukaryotes. Some clues have been provided from work performed in yeast showing that ATM and PIF1 regulate DNTF. ATM and ATR have overlapping functions in Arabidopsis, making each of them potential candidates for regulation of DNTF. There are three putative PIF1 proteins in Arabidopsis, each roughly a third of the size of ScPIF1. In addition, two of these genes are right next to one another, making genetic analysis of AtPIF1 prohibitive.

As mentioned above, a putative complex consisting of Ku, POT1b, and TER2 may be important for protecting chromosome ends and regulating DNTF. Analysis is currently underway to determine DNTF efficiencies in POT1b and TER2 mutants. Shorter TRAs (<100bp) are being tested alongside longer (~2kb) TRAs in order to more closely mimic DNTF at a DSB *in vivo*.

The role of Ku in DNTF can be also be probed further. A TER binding domain is conserved between all Ku80 molecules that are known to bind RNA. Therefore, it would be interesting to investigate how mutating this RNA binding motif affects DNTF. An important control for this experiment would be determining what effect this mutation has on DSB repair and telomere length maintenance. The RNA binding domain of Ku80 may be critical for recruitment of TER2 and POT1b to telomeres, and therefore mimic the *ter2pot1b* background described in Chapter IV.

To perform this experiment, Ku80 RNA binding mutants ($Ku80_{RBM}$) would be used to complement a *Ku80* background. Complementation would be judged by the criteria listed above and compared to a complementation with wild type Ku80. If the

two-faced Ku model applies to Arabidopsis, then $Ku80_{RBM}$ should not have an effect on NHEJ, but might be defective for telomere length maintenance (Ribes-Zamora, 2007).

ATR is a fascinating kinase for many reasons, but its large size has limited its biochemical characterization. ATR regulates many activities at both DSBs and the chromosome end, and is an excellent candidate for regulation of DNTF. In Arabidopsis, ATR is also important for telomerase regulation, as plants lacking ATR for at least two generations show a significant reduction in telomerase activity (Appendix A-3). Despite this, telomeres are maintained at a normal length in these lines over multiple generations (Vespa et al 2005). This suggests that ATR is not critical for maintenance of normal telomeres.

However, de novo telomeres are not wild type telomeres initially. Due to nuances in how TRAs are inserted into the genome, a large degree of degradation (up to ~50% of the TRA) can occur. In a wild type genetic background, this is a stochastic process and mechanisms are in place to protect and extend the new telomere. In contrast, 4X *atr* mutants show increased degradation of the TRA prior to stabilization, providing a clue for ATR's role in telomere biology in Arabidopsis.

Transformation of 4X atr lines with a 2kb TRA resulted in a DNTF efficiency only slightly lower than WT (40% vs 56%, n=8/20). Strikingly, 6/8 DNTF lines established *de novo* telomeres that were much shorter than wild type (~1kb). PETRA analysis of the progeny revealed telomeres of the same length. Thus, the *de novo* telomeres were being maintained, but were incapable of extending into the wild type range.

These data suggest that ATR is necessary for the recognition of short telomeres, or the recruitment of telomerase to short telomeres in order to extend them to a normal length. This is similar to that observed in yeast, but contrary to the situation in vertebrates. What happens if short TRAs are transformed into the 4x *atr* setting? In Arabidopsis, telomeres below 1kb are prone to activating DDR pathways and become caught up in chromosomal fusions (Heacock et al, 2007). Is there a secondary regulatory mechanism to return the short TRAs past the 1kb threshold in order to avoid telomere fusions?

Conclusions

In summary, this dissertation has highlighted the diverse mechanisms by which telomerase is regulated in Arabidopsis. These mechanisms include inhibition by an alternative telomerase RNA, TER2, that is activated during DNA damage. In the process of analyzing TER2 I discovered what could be a novel family of transposable elements conserved in the Brassicaceae lineage. I characterized POT1b and POT1c and found that they perform complex roles at the chromosome end. I also determined several factors that are important in the regulation of telomere addition at DSBs. These findings underscore the diversity of mechanisms by which the cell defines telomeres versus DSBs.

REFERENCES

Amiard S, Charbonnel C, Allain E, Depeiges A, White CI, Gallego ME (2010) Distinct roles of the ATR kinase and the Mre11-Rad50-Nbs1 complex in the maintenance of chromosomal stability in Arabidopsis. *Plant Cell* **22:** 3020-3033

Amiard S, Depeiges A, Allain E, White CI, Gallego ME (2011) Arabidopsis ATM and ATR kinases prevent propagation of genome damage caused by telomere dysfunction. *Plant Cell* **23:** 4254-4265

Amiard S, Doudeau M, Pinte S, Poulet A, Lenain C, Faivre-Moskalenko C, Angelov D, Hug N, Vindigni A, Bouvet P, Paoletti J, Gilson E, Giraud-Panis MJ (2007) A topological mechanism for TRF2-enhanced strand invasion. *Nat Struct Mol Biol* **14:** 147-154

Anderson BH, Kasher PR, Mayer J, Szynkiewicz M, Jenkinson EM, Bhaskar SS, Urquhart JE, Daly SB, Dickerson JE, O'Sullivan J, Leibundgut EO, Muter J, Abdel-Salem GM, Babul-Hirji R, Baxter P, Berger A, Bonafe L, Brunstom-Hernandez JE, Buckard JA, Chitayat D, Chong WK, Cordelli DM, Ferreira P, Fluss J, Forrest EH, Franzoni E, Garone C, Hammans SR, Houge G, Hughes I, Jacquemont S, Jeannet PY, Jefferson RJ, Kumar R, Kutschke G, Lundberg S, Lourenco CM, Mehta R, Naidu S, Nischal KK, Nunes L, Ounap K, Philippart M, Prabhakar P, Risen SR, Schiffmann R, Soh C, Stephenson JB, Stewart H, Stone J, Tolmie JL, van der Knaap MS, Vieira JP, Vilain CN, Wakeling EL, Wermenbol V, Whitney A, Lovell SC, Meyer S, Livingston JH, Baerlocher GM, Black GC, Rice GI, Crow YJ (2012) Mutations in CTC1, encoding conserved telomere maintenance component 1, cause Coats plus. *Nat Genet* **44:** 338-342

Arneric M, Lingner J (2007) Tel1 kinase and subtelomere-bound Tbf1 mediate preferential elongation of short telomeres by telomerase in yeast. *EMBO Rep* **8:** 1080-1085

Audebert M, Salles B, Calsou P (2004) Involvement of poly(ADP-ribose) polymerase-1 and XRCC1/DNA ligase III in an alternative route for DNA double-strand breaks rejoining. *J Biol Chem* **279:** 55117-55126

Autexier C, Lue NF (2006) The structure and function of telomerase reverse transcriptase. *Annu Rev Biochem* **75:** 493-517

Bae NS, Baumann P (2007) A RAP1/TRF2 complex inhibits nonhomologous end-joining at human telomeric DNA ends. *Mol Cell* **26:** 323-334

Barnett MA, Buckle VJ, Evans EP, Porter AC, Rout D, Smith AG, Brown WR (1993) Telomere directed fragmentation of mammalian chromosomes. *Nucleic Acids Res* **21:** 27-36

Batista LF, Pech MF, Zhong FL, Nguyen HN, Xie KT, Zaug AJ, Crary SM, Choi J, Sebastiano V, Cherry A, Giri N, Wernig M, Alter BP, Cech TR, Savage SA, Reijo Pera RA, Artandi SE (2011) Telomere shortening and loss of self-renewal in dyskeratosis congenita induced pluripotent stem cells. *Nature* **474:** 399-402

Baumann P, Cech TR (2001) Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science* **292:** 1171-1175

Baumann P, Podell E, Cech TR (2002) Human Pot1 (protection of telomeres) protein: cytolocalization, gene structure, and alternative splicing. *Mol Cell Biol* **22:** 8079-8087

Beernink HT, Miller K, Deshpande A, Bucher P, Cooper JP (2003) Telomere maintenance in fission yeast requires an Est1 ortholog. *Curr Biol* **13:** 575-580

Beilstein MA, Nagalingum NS, Clements MD, Manchester SR, Mathews S (2010) Dated molecular phylogenies indicate a Miocene origin for Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **107:** 18724-18728

Ben Amor B, Wirth S, Merchan F, Laporte P, d'Aubenton-Carafa Y, Hirsch J, Maizel A, Mallory A, Lucas A, Deragon JM, Vaucheret H, Thermes C, Crespi M (2009) Novel long non-protein coding RNAs involved in Arabidopsis differentiation and stress responses. *Genome Res* **19:** 57-69

Bertuch AA, Lundblad V (2003) The Ku heterodimer performs separable activities at double-strand breaks and chromosome termini. *Mol Cell Biol* **23:** 8202-8215

Bertuch AA, Lundblad V (2003) Which end: dissecting Ku's function at telomeres and double-strand breaks. *Genes Dev* **17:** 2347-2350

Bertuch AA, Lundblad V (2006) The maintenance and masking of chromosome termini. *Curr Opin Cell Biol* **18:** 247-253

Bianchi A, Negrini S, Shore D (2004) Delivery of yeast telomerase to a DNA break depends on the recruitment functions of Cdc13 and Est1. *Mol Cell* **16:** 139-146

Bianchi A, Shore D (2007) Early replication of short telomeres in budding yeast. *Cell* **128:** 1051-1062

Bianchi A, Shore D (2008) How telomerase reaches its end: mechanism of telomerase regulation by the telomeric complex. *Mol Cell* **31:** 153-165

Bianchi A, Smith S, Chong L, Elias P, de Lange T (1997) TRF1 is a dimer and bends telomeric DNA. *EMBO J* **16:** 1785-1794

Bilaud T, Brun C, Ancelin K, Koering CE, Laroche T, Gilson E (1997) Telomeric localization of TRF2, a novel human telobox protein. *Nat Genet* **17:** 236-239

Bilaud T, Koering CE, Binet-Brasselet E, Ancelin K, Pollice A, Gasser SM, Gilson E (1996) The telobox, a Myb-related telomeric DNA binding motif found in proteins from yeast, plants and human. *Nucleic Acids Res* **24:** 1294-1303

Birchler JA, Yu W, Han F (2008) Plant engineered minichromosomes and artificial chromosome platforms. *Cytogenet Genome Res* **120:** 228-232

Blackburn EH, Gall JG (1978) A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena. *J Mol Biol* **120:** 33-53

Bochkarev A, Bochkareva E (2004) From RPA to BRCA2: lessons from single-stranded DNA binding by the OB-fold. *Curr Opin Struct Biol* **14:** 36-42

Boltz KA, Leehy K, Song X, Nelson AD, Shippen DE (2012) ATR cooperates with CTC1 and STN1 to maintain telomeres and genome integrity in Arabidopsis. *Mol Biol Cell* **23:** 1558-1568

Bonaglia MC, Giorda R, Beri S, De Agostini C, Novara F, Fichera M, Grillo L, Galesi O, Vetro A, Ciccone R, Bonati MT, Giglio S, Guerrini R, Osimani S, Marelli S, Zucca C, Grasso R, Borgatti R, Mani E, Motta C, Molteni M, Romano C, Greco D, Reitano S, Baroncini A, Lapi E, Cecconi A, Arrigo G, Patricelli MG, Pantaleoni C, D'Arrigo S, Riva D, Sciacca F, Dalla Bernardina B, Zoccante L, Darra F, Termine C, Maserati E, Bigoni S, Priolo E, Bottani A, Gimelli S, Bena F, Brusco A, di Gregorio E, Bagnasco I, Giussani U, Nitsch L, Politi P, Martinez-Frias ML, Martinez-Fernandez ML, Martinez Guardia N, Bremer A, Anderlid BM, Zuffardi O (2011) Molecular mechanisms generating and stabilizing terminal 22q13 deletions in 44 subjects with Phelan/McDermid syndrome. *PLoS Genet* **7:** e1002173

Bonen L (2008) Cis- and trans-splicing of group II introns in plant mitochondria. *Mitochondrion* **8:** 26-34

Bonetti D, Clerici M, Manfrini N, Lucchini G, Longhese MP (2010) The MRX complex plays multiple functions in resection of Yku- and Rif2-protected DNA ends. *PLoS One* **5:** e14142

Bosoy D, Peng Y, Mian IS, Lue NF (2003) Conserved N-terminal motifs of telomerase reverse transcriptase required for ribonucleoprotein assembly in vivo. *J Biol Chem* **278:** 3882-3890

Boule JB, Vega LR, Zakian VA (2005) The yeast Pif1p helicase removes telomerase from telomeric DNA. *Nature* **438:** 57-61

Boulton SJ, Jackson SP (1996) Identification of a Saccharomyces cerevisiae Ku80 homologue: roles in DNA double strand break rejoining and in telomeric maintenance. *Nucleic Acids Res* **24:** 4639-4648

Boulton SJ, Jackson SP (1996) Saccharomyces cerevisiae Ku70 potentiates illegitimate DNA double-strand break repair and serves as a barrier to error-prone DNA repair pathways. *EMBO J* **15:** 5093-5103

Bourns BD, Alexander MK, Smith AM, Zakian VA (1998) Sir proteins, Rif proteins, and Cdc13p bind Saccharomyces telomeres in vivo. *Mol Cell Biol* **18:** 5600-5608

Box JA, Bunch JT, Zappulla DC, Glynn EF, Baumann P (2008) A flexible template boundary element in the RNA subunit of fission yeast telomerase. *J Biol Chem* **283:** 24224-24233

Britt AB, May GD (2003) Re-engineering plant gene targeting. *Trends Plant Sci* **8:** 90-95

Broccoli D, Smogorzewska A, Chong L, de Lange T (1997) Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. *Nat Genet* **17:** 231-235

Brown EJ, Baltimore D (2000) ATR disruption leads to chromosomal fragmentation and early embryonic lethality. *Genes Dev* **14:** 397-402

Bryan TM, Marusic L, Bacchetti S, Namba M, Reddel RR (1997) The telomere lengthening mechanism in telomerase-negative immortal human cells does not involve the telomerase RNA subunit. *Hum Mol Genet* **6:** 921-926

Bundock P, van Attikum H, Hooykaas P (2002) Increased telomere length and hypersensitivity to DNA damaging agents in an *Arabidopsis* KU70 mutant. *Nucleic Acids Res* **30:** 3395-3400

Bureau TE, Wessler SR (1992) Tourist: a large family of small inverted repeat elements frequently associated with maize genes. *Plant Cell* **4:** 1283-1294

Bycroft M, Hubbard TJ, Proctor M, Freund SM, Murzin AG (1997) The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. *Cell* **88:** 235-242

Cao J, Schneeberger K, Ossowski S, Gunther T, Bender S, Fitz J, Koenig D, Lanz C, Stegle O, Lippert C, Wang X, Ott F, Muller J, Alonso-Blanco C, Borgwardt K, Schmid KJ, Weigel D (2011) Whole-genome sequencing of multiple Arabidopsis thaliana populations. *Nat Genet* **43:** 956-963

Carvalho A, Delgado M, Barao A, Frescatada M, Ribeiro E, Pikaard CS, Viegas W, Neves N (2010) Chromosome and DNA methylation dynamics during meiosis in the autotetraploid Arabidopsis arenosa. *Sex Plant Reprod* **23:** 29-37

Casteel DE, Zhuang S, Zeng Y, Perrino FW, Boss GR, Goulian M, Pilz RB (2009) A DNA polymerase-{alpha}{middle dot}primase cofactor with homology to replication protein A-32 regulates DNA replication in mammalian cells. *J Biol Chem* **284:** 5807-5818

Celli GB, de Lange T (2005) DNA processing is not required for ATM-mediated telomere damage response after TRF2 deletion. *Nat Cell Biol* **7:** 712-718

Cesare AJ, Groff-Vindman C, Compton SA, McEachern MJ, Griffith JD (2008) Telomere loops and homologous recombination-dependent telomeric circles in a Kluyveromyces lactis telomere mutant strain. *Mol Cell Biol* **28:** 20-29

Cesare AJ, Quinney N, Willcox S, Subramanian D, Griffith JD (2003) Telomere looping in P. sativum (common garden pea). *Plant J* **36:** 271-279

Chang M, Arneric M, Lingner J (2007) Telomerase repeat addition processivity is increased at critically short telomeres in a Tel1-dependent manner in Saccharomyces cerevisiae. *Genes Dev* **21:** 2485-2494

Chapon C, Cech TR, Zaug AJ (1997) Polyadenylation of telomerase RNA in budding yeast. *RNA* **3:** 1337-1351

Charon C, Moreno AB, Bardou F, Crespi M (2010) Non-protein-coding RNAs and their interacting RNA-binding proteins in the plant cell nucleus. *Mol Plant* **3:** 729-739

Chawla R, Redon S, Raftopoulou C, Wischnewski H, Gagos S, Azzalin CM (2011) Human UPF1 interacts with TPP1 and telomerase and sustains telomere leading-strand replication. *EMBO J* **30:** 4047-4058

Chen JL, Greider CW (2004) Telomerase RNA structure and function: implications for dyskeratosis congenita. *Trends Biochem Sci* **29:** 183-192

Chen JL, Greider CW (2005) Functional analysis of the pseudoknot structure in human telomerase RNA. *Proc Natl Acad Sci U S A* **102:** 8080-8085; discussion 8077-8089

Cheng F, Liu S, Wu J, Fang L, Sun S, Liu B, Li P, Hua W, Wang X (2011) BRAD, the genetics and genomics database for Brassica plants. *BMC Plant Biol* **11:** 136

Chiu SY, Serin G, Ohara O, Maquat LE (2003) Characterization of human Smg5/7a: a protein with similarities to Caenorhabditis elegans SMG5 and SMG7 that functions in the dephosphorylation of Upf1. *RNA* **9:** 77-87

Chong L, van Steensel B, Broccoli D, Erdjument-Bromage H, Hanish J, Tempst P, de Lange T (1995) A human telomeric protein. *Science* **270:** 1663-1667

Chung WH, Zhu Z, Papusha A, Malkova A, Ira G (2010) Defective resection at DNA double-strand breaks leads to de novo telomere formation and enhances gene targeting. *PLoS Genet* **6:** e1000948

Churikov D, Price CM (2008) Pot1 and cell cycle progression cooperate in telomere length regulation. *Nat Struct Mol Biol* **15:** 79-84

Churikov D, Wei C, Price CM (2006) Vertebrate POT1 restricts G-overhang length and prevents activation of a telomeric DNA damage checkpoint but is dispensable for overhang protection. *Mol Cell Biol* **26:** 6971-6982

Cifuentes-Rojas C, Kannan K, Tseng L, Shippen DE (2011) Two RNA subunits and POT1a are components of Arabidopsis telomerase. *Proc Natl Acad Sci U S A* **108:** 73-78

Cifuentes-Rojas C, Shippen DE (2012) Telomerase regulation. *Mutat Res* **730:** 20-27

Clark MB, Johnston RL, Inostroza-Ponta M, Fox AH, Fortini E, Moscato P, Dinger ME, Mattick JS (2012) Genome-wide analysis of long noncoding RNA stability. *Genome Res*

Cohen SB, Graham ME, Lovrecz GO, Bache N, Robinson PJ, Reddel RR (2007) Protein composition of catalytically active human telomerase from immortal cells. *Science* **315:** 1850-1853

Colgin LM, Baran K, Baumann P, Cech TR, Reddel RR (2003) Human POT1 facilitates telomere elongation by telomerase. *Curr Biol* **13:** 942-946

Collins K (2006) The biogenesis and regulation of telomerase holoenzymes. *Nat Rev Mol Cell Biol* **7:** 484-494

Conrad MN, Wright JH, Wolf AJ, Zakian VA (1990) RAP1 protein interacts with yeast telomeres in vivo: overproduction alters telomere structure and decreases chromosome stability. *Cell* **63:** 739-750

Counter CM, Meyerson M, Eaton EN, Weinberg RA (1997) The catalytic subunit of yeast telomerase. *Proc Natl Acad Sci U S A* **94:** 9202-9207

Cove D, Bezanilla M, Harries P, Quatrano R (2006) Mosses as model systems for the study of metabolism and development. *Annu Rev Plant Biol* **57:** 497-520

Crabbe L, Verdun RE, Haggblom CI, Karlseder J (2004) Defective telomere lagging strand synthesis in cells lacking WRN helicase activity. *Science* **306:** 1951-1953

Craven RJ, Petes TD (1999) Dependence of the regulation of telomere length on the type of subtelomeric repeat in the yeast Saccharomyces cerevisiae. *Genetics* **152:** 1531-1541

Croy JE, Wuttke DS (2006) Themes in ssDNA recognition by telomere-end protection proteins. *Trends Biochem Sci* **31:** 516-525

Culligan K, Tissier A, Britt A (2004) ATR regulates a G2-phase cell-cycle checkpoint in Arabidopsis thaliana. *Plant Cell* **16:** 1091-1104

Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible WR (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiol* **139:** 5-17

Dai L, Toor N, Olson R, Keeping A, Zimmerly S (2003) Database for mobile group II introns. *Nucleic Acids Res* **31:** 424-426

de Bruin D, Kantrow SM, Liberatore RA, Zakian VA (2000) Telomere folding is required for the stable maintenance of telomere position effects in yeast. *Mol Cell Biol* **20:** 7991-8000

de Klein A, Muijtjens M, van Os R, Verhoeven Y, Smit B, Carr AM, Lehmann AR, Hoeijmakers JH (2000) Targeted disruption of the cell-cycle checkpoint gene ATR leads to early embryonic lethality in mice. *Curr Biol* **10:** 479-482

de Lange T (2005) Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev* **19:** 2100-2110

de Lange T (2009) How telomeres solve the end-protection problem. *Science* **326:** 948-952

Denchi EL (2009) Give me a break: how telomeres suppress the DNA damage response. *DNA Repair (Amst)* **8:** 1118-1126

Denchi EL, de Lange T (2007) Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. *Nature* **448:** 1068-1071

Deng Y, Guo X, Ferguson DO, Chang S (2009) Multiple roles for MRE11 at uncapped telomeres. *Nature* **460:** 914-918

Diede SJ, Gottschling DE (1999) Telomerase-mediated telomere addition in vivo requires DNA primase and DNA polymerases alpha and delta. *Cell* **99:** 723-733

Diede SJ, Gottschling DE (2001) Exonuclease activity is required for sequence addition and Cdc13p loading at a de novo telomere. *Curr Biol* **11:** 1336-1340

Dionne I, Wellinger RJ (1998) Processing of telomeric DNA ends requires the passage of a replication fork. *Nucleic Acids Res* **26:** 5365-5371

Dissmeyer N, Weimer AK, Pusch S, De Schutter K, Alvim Kamei CL, Nowack MK, Novak B, Duan GL, Zhu YG, De Veylder L, Schnittger A (2009) Control of cell proliferation, organ growth, and DNA damage response operate independently of dephosphorylation of the Arabidopsis Cdk1 homolog CDKA;1. *Plant Cell* **21:** 3641-3654

Doucet-Chabeaud G, Godon C, Brutesco C, de Murcia G, Kazmaier M (2001) Ionising radiation induces the expression of PARP-1 and PARP-2 genes in Arabidopsis. *Mol Genet Genomics* **265:** 954-963

DuBois ML, Haimberger ZW, McIntosh MW, Gottschling DE (2002) A quantitative assay for telomere protection in *Saccharomyces cerevisiae*. *Genetics* **161:** 995-1013

Eugster A, Lanzuolo C, Bonneton M, Luciano P, Pollice A, Pulitzer JF, Stegberg E, Berthiau AS, Forstemann K, Corda Y, Lingner J, Geli V, Gilson E (2006) The finger subdomain of yeast telomerase cooperates with Pif1p to limit telomere elongation. *Nat Struct Mol Biol* **13:** 734-739

Evans SK, Lundblad V (2002) The Est1 subunit of Saccharomyces cerevisiae telomerase makes multiple contributions to telomere length maintenance. *Genetics* **162:** 1101-1115

Fajkus J, Kovarik A, Kralovics R, Bezdek M (1995) Organization of telomeric and subtelomeric chromatin in the higher plant Nicotiana tabacum. *Mol Gen Genet* **247:** 633-638

Fattah F, Lee EH, Weisensel N, Wang Y, Lichter N, Hendrickson EA (2010) Ku regulates the non-homologous end joining pathway choice of DNA double-strand break repair in human somatic cells. *PLoS Genet* **6:** e1000855

Feng J, Funk WD, Wang SS, Weinrich SL, Avilion AA, Chiu CP, Adams RR, Chang E, Allsopp RC, Yu J, et al. (1995) The RNA component of human telomerase. *Science* **269:** 1236-1241

Ferreira MG, Miller KM, Cooper JP (2004) Indecent exposure: when telomeres become uncapped. *Mol Cell* **13:** 7-18

Feschotte C, Jiang N, Wessler SR (2002) Plant transposable elements: where genetics meets genomics. *Nat Rev Genet* **3:** 329-341

Feschotte C, Pritham EJ (2007) DNA transposons and the evolution of eukaryotic genomes. *Annu Rev Genet* **41:** 331-368

Fisher TS, Taggart AK, Zakian VA (2004) Cell cycle-dependent regulation of yeast telomerase by Ku. *Nat Struct Mol Biol* **11:** 1198-1205

Fisher TS, Zakian VA (2005) Ku: a multifunctional protein involved in telomere maintenance. *DNA Repair (Amst)* **4:** 1215-1226

Fitzgerald MS, Riha K, Gao F, Ren S, McKnight TD, Shippen DE (1999) Disruption of the telomerase catalytic subunit gene from Arabidopsis inactivates telomerase and leads to a slow loss of telomeric DNA. *Proc Natl Acad Sci U S A* **96:** 14813-14818

Fitzgerald MS, Shakirov EV, Hood EE, McKnight TD, Shippen DE (2001) Different modes of de novo telomere formation by plant telomerases. *Plant J* **26:** 77-87

Flint J, Craddock CF, Villegas A, Bentley DP, Williams HJ, Galanello R, Cao A, Wood WG, Ayyub H, Higgs DR (1994) Healing of broken human chromosomes by the addition of telomeric repeats. *Am J Hum Genet* **55:** 505-512

Flynn RL, Centore RC, O'Sullivan RJ, Rai R, Tse A, Songyang Z, Chang S, Karlseder J, Zou L (2011) TERRA and hnRNPA1 orchestrate an RPA-to-POT1 switch on telomeric single-stranded DNA. *Nature* **471:** 532-536

Flynn RL, Chang S, Zou L (2012) RPA and POT1: friends or foes at telomeres? *Cell Cycle* **11:** 652-657

Forsyth NR, Wright WE, Shay JW (2002) Telomerase and differentiation in multicellular organisms: turn it off, turn it on, and turn it off again. *Differentiation* **69:** 188-197

Fortin F, Beaulieu Bergeron M, Fetni R, Lemieux N (2009) Frequency of chromosome healing and interstitial telomeres in 40 cases of constitutional abnormalities. *Cytogenet Genome Res* **125:** 176-185

Fouche N, Cesare AJ, Willcox S, Ozgur S, Compton SA, Griffith JD (2006) The basic domain of TRF2 directs binding to DNA junctions irrespective of the presence of TTAGGG repeats. *J Biol Chem* **281:** 37486-37495

Friesner J, Britt AB (2003) Ku80- and DNA ligase IV-deficient plants are sensitive to ionizing radiation and defective in T-DNA integration. *Plant J* **34:** 427-440

Fu D, Collins K (2003) Distinct biogenesis pathways for human telomerase RNA and H/ACA small nucleolar RNAs. *Mol Cell* **11:** 1361-1372

Fulcher N, Sablowski R (2009) Hypersensitivity to DNA damage in plant stem cell niches. *Proc Natl Acad Sci U S A* **106:** 20984-20988

Furukawa T, Curtis MJ, Tominey CM, Duong YH, Wilcox BW, Aggoune D, Hays JB, Britt AB (2010) A shared DNA-damage-response pathway for induction of stem-cell death by UVB and by gamma irradiation. *DNA Repair (Amst)* **9:** 940-948

Gallardo F, Laterreur N, Cusanelli E, Ouenzar F, Querido E, Wellinger RJ, Chartrand P (2011) Live cell imaging of telomerase RNA dynamics reveals cell cycle-dependent clustering of telomerase at elongating telomeres. *Mol Cell* **44:** 819-827

Gallego ME, Bleuyard JY, Daoudal-Cotterell S, Jallut N, White CI (2003) Ku80 plays a role in non-homologous recombination but is not required for T-DNA integration in *Arabidopsis*. *Plant J* **35:** 557-565

Gallego ME, Jalut N, White CI (2003) Telomerase dependence of telomere lengthening in Ku80 mutant Arabidopsis. *Plant Cell* **15:** 782-789

Gao Q, Reynolds GE, Wilcox A, Miller D, Cheung P, Artandi SE, Murnane JP (2008) Telomerase-dependent and -independent chromosome healing in mouse embryonic stem cells. *DNA Repair (Amst)* **7:** 1233-1249

Garcia V, Bruchet H, Camescasse D, Granier F, Bouchez D, Tissier A (2003) AtATM is essential for meiosis and the somatic response to DNA damage in plants. *Plant Cell* **15:** 119-132

Garvik B, Carson M, Hartwell L (1995) Single-stranded DNA arising at telomeres in cdc13 mutants may constitute a specific signal for the RAD9 checkpoint. *Mol Cell Biol* **15:** 6128-6138

Gelvin SB (2003) Agrobacterium-mediated plant transformation: the biology behind the "gene-jockeying" tool. *Microbiol Mol Biol Rev* **67:** 16-37, table of contents

Giraud-Panis MJ, Pisano S, Poulet A, Le Du MH, Gilson E (2010) Structural identity of telomeric complexes. *FEBS Lett* **584:** 3785-3799

Gladyshev EA, Arkhipova IR (2007) Telomere-associated endonuclease-deficient Penelope-like retroelements in diverse eukaryotes. *Proc Natl Acad Sci U S A* **104:** 9352-9357

Gong Y, de Lange T (2010) A Shld1-controlled POT1a provides support for repression of ATR signaling at telomeres through RPA exclusion. *Mol Cell* **40:** 377-387

Gottschling DE, Zakian VA (1986) Telomere proteins: specific recognition and protection of the natural termini of Oxytricha macronuclear DNA. *Cell* **47:** 195-205

Grandin N, Damon C, Charbonneau M (2001) Ten1 functions in telomere end protection and length regulation in association with Stn1 and Cdc13. *EMBO J* **20:** 1173-1183

Grandin N, Reed SI, Charbonneau M (1997) Stn1, a new Saccharomyces cerevisiae protein, is implicated in telomere size regulation in association with Cdc13. *Genes Dev* **11:** 512-527

Gray JT, Celander DW, Price CM, Cech TR (1991) Cloning and expression of genes for the Oxytricha telomere-binding protein: specific subunit interactions in the telomeric complex. *Cell* **67:** 807-814

Greider CW, Blackburn EH (1985) Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. *Cell* **43:** 405-413

Greider CW, Blackburn EH (1987) The telomere terminal transferase of Tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell* **51:** 887-898

Greider CW, Blackburn EH (1989) A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis. *Nature* **337:** 331-337

Griffith JD, Comeau L, Rosenfield S, Stansel RM, Bianchi A, Moss H, de Lange T (1999) Mammalian telomeres end in a large duplex loop. *Cell* **97:** 503-514

Grossi S, Bianchi A, Damay P, Shore D (2001) Telomere formation by rap1p binding site arrays reveals end-specific length regulation requirements and active telomeric recombination. *Mol Cell Biol* **21:** 8117-8128

Haendeler J, Hoffmann J, Diehl JF, Vasa M, Spyridopoulos I, Zeiher AM, Dimmeler S (2004) Antioxidants inhibit nuclear export of telomerase reverse transcriptase and delay replicative senescence of endothelial cells. *Circ Res* **94:** 768-775

Hanish JP, Yanowitz JL, de Lange T (1994) Stringent sequence requirements for the formation of human telomeres. *Proc Natl Acad Sci U S A* **91:** 8861-8865

Hannes F, Van Houdt J, Quarrell OW, Poot M, Hochstenbach R, Fryns JP, Vermeesch JR (2010) Telomere healing following DNA polymerase arrest-induced breakages is likely the main mechanism generating chromosome 4p terminal deletions. *Hum Mutat* **31:** 1343-1351

Hardy CF, Sussel L, Shore D (1992) A RAP1-interacting protein involved in transcriptional silencing and telomere length regulation. *Genes Dev* **6:** 801-814

Harries PA, Pan A, Quatrano RS (2005) Actin-related protein2/3 complex component ARPC1 is required for proper cell morphogenesis and polarized cell growth in Physcomitrella patens. *Plant Cell* **17:** 2327-2339

Harrington L, Zhou W, McPhail T, Oulton R, Yeung DS, Mar V, Bass MB, Robinson MO (1997) Human telomerase contains evolutionarily conserved catalytic and structural subunits. *Genes Dev* **11:** 3109-3115

Hashimura Y, Ueguchi C (2011) The Arabidopsis MERISTEM DISORGANIZATION 1 gene is required for the maintenance of stem cells through the reduction of DNA damage. *Plant J* **68:** 657-669

Hayflick L, Moorhead PS (1961) The serial cultivation of human diploid cell strains. *Exp Cell Res* **25:** 585-621

He H, Multani AS, Cosme-Blanco W, Tahara H, Ma J, Pathak S, Deng Y, Chang S (2006) POT1b protects telomeres from end-to-end chromosomal fusions and aberrant homologous recombination. *EMBO J* **25:** 5180-5190

Heacock M, Spangler E, Riha K, Puizina J, Shippen DE (2004) Molecular analysis of telomere fusions in Arabidopsis: multiple pathways for chromosome end-joining. *EMBO J* **23:** 2304-2313

Heacock ML, Idol RA, Friesner JD, Britt AB, Shippen DE (2007) Telomere dynamics and fusion of critically shortened telomeres in plants lacking DNA ligase IV. *Nucleic Acids Res* **35:** 6490-6500

Hefferin ML, Tomkinson AE (2005) Mechanism of DNA double-strand break repair by non-homologous end joining. *DNA Repair (Amst)* **4:** 639-648

Hefner E, Huefner N, Britt AB (2006) Tissue-specific regulation of cell-cycle responses to DNA damage in Arabidopsis seedlings. *DNA Repair (Amst)* **5:** 102-110

Henderson ER, Blackburn EH (1989) An overhanging 3' terminus is a conserved feature of telomeres. *Mol Cell Biol* **9:** 345-348

Henry IM, Dilkes BP, Young K, Watson B, Wu H, Comai L (2005) Aneuploidy and genetic variation in the *Arabidopsis thaliana* triploid response. *Genetics* **170:** 1979-1988

Higashiyama T, Noutoshi Y, Akiba M, Yamada T (1995) Telomere and LINE-like elements at the termini of the Chlorella chromosome I. *Nucleic Acids Symp Ser***:** 71-72

Hirano Y, Sugimoto K (2007) Cdc13 telomere capping decreases Mec1 association but does not affect Tel1 association with DNA ends. *Mol Biol Cell* **18:** 2026-2036

Hockemeyer D, Daniels JP, Takai H, de Lange T (2006) Recent expansion of the telomeric complex in rodents: Two distinct POT1 proteins protect mouse telomeres. *Cell* **126:** 63-77

Hockemeyer D, Palm W, Else T, Daniels JP, Takai KK, Ye JZ, Keegan CE, de Lange T, Hammer GD (2007) Telomere protection by mammalian Pot1 requires interaction with Tpp1. *Nat Struct Mol Biol* **14:** 754-761

Hoen DR, Park KC, Elrouby N, Yu Z, Mohabir N, Cowan RK, Bureau TE (2006) Transposon-mediated expansion and diversification of a family of ULP-like genes. *Mol Biol Evol* **23:** 1254-1268

Hong JP, Byun MY, Koo DH, An K, Bang JW, Chung IK, An G, Kim WT (2007) Suppression of RICE TELOMERE BINDING PROTEIN 1 results in severe and gradual developmental defects accompanied by genome instability in rice. *Plant Cell* **19:** 1770-1781

Houghtaling BR, Cuttonaro L, Chang W, Smith S (2004) A dynamic molecular link between the telomere length regulator TRF1 and the chromosome end protector TRF2. *Curr Biol* **14:** 1621-1631

Hsu M, McEachern MJ, Dandjinou AT, Tzfati Y, Orr E, Blackburn EH, Lue NF (2007) Telomerase core components protect Candida telomeres from aberrant overhang accumulation. *Proc Natl Acad Sci U S A* **104:** 11682-11687

Hu TT, Pattyn P, Bakker EG, Cao J, Cheng JF, Clark RM, Fahlgren N, Fawcett JA, Grimwood J, Gundlach H, Haberer G, Hollister JD, Ossowski S, Ottilar RP, Salamov AA, Schneeberger K, Spannagl M, Wang X, Yang L, Nasrallah ME, Bergelson J, Carrington JC, Gaut BS, Schmutz J, Mayer KF, Van de Peer Y, Grigoriev IV, Nordborg M, Weigel D, Guo YL (2011) The Arabidopsis lyrata genome sequence and the basis of rapid genome size change. *Nat Genet* **43:** 476-481

Hug N, Lingner J (2006) Telomere length homeostasis. *Chromosoma* **115:** 413-425

Itzhaki JE, Barnett MA, MacCarthy AB, Buckle VJ, Brown WR, Porter AC (1992) Targeted breakage of a human chromosome mediated by cloned human telomeric DNA. *Nat Genet* **2:** 283-287

Jacob NK, Lescasse R, Linger BR, Price CM (2007) Tetrahymena POT1a regulates telomere length and prevents activation of a cell cycle checkpoint. *Mol Cell Biol* **27:** 1592-1601

Jacob NK, Skopp R, Price CM (2001) G-overhang dynamics at Tetrahymena telomeres. *EMBO J* **20:** 4299-4308

Jain D, Cooper JP (2010) Telomeric strategies: means to an end. *Annu Rev Genet* **44:** 243-269

Jamonnak N, Creamer TJ, Darby MM, Schaughency P, Wheelan SJ, Corden JL (2011) Yeast Nrd1, Nab3, and Sen1 transcriptome-wide binding maps suggest multiple roles in post-transcriptional RNA processing. *RNA* **17:** 2011-2025

Jiang N, Bao Z, Zhang X, Eddy SR, Wessler SR (2004) Pack-MULE transposable elements mediate gene evolution in plants. *Nature* **431:** 569-573

Jiang N, Bao Z, Zhang X, Hirochika H, Eddy SR, McCouch SR, Wessler SR (2003) An active DNA transposon family in rice. *Nature* **421:** 163-167

Johnston JS, Pepper AE, Hall AE, Chen ZJ, Hodnett G, Drabek J, Lopez R, Price HJ (2005) Evolution of genome size in Brassicaceae. *Ann Bot* **95:** 229-235

Kannan K, Nelson AD, Shippen DE (2008) Dyskerin is a component of the Arabidopsis telomerase RNP required for telomere maintenance. *Mol Cell Biol* **28:** 2332-2341

Kanoh J, Ishikawa F (2003) Composition and conservation of the telomeric complex. *Cell Mol Life Sci* **60:** 2295-2302

Karamysheva ZN, Surovtseva YV, Vespa L, Shakirov EV, Shippen DE (2004) A C-terminal Myb extension domain defines a novel family of double-strand telomeric DNA-binding proteins in Arabidopsis. *J Biol Chem* **279:** 47799-47807

Karimi M, Inze D, Depicker A (2002) GATEWAY vectors for *Agrobacterium*-mediated plant transformation. *Trends Plant Sci* **7:** 193-195

Keren I, Bezawork-Geleta A, Kolton M, Maayan I, Belausov E, Levy M, Mett A, Gidoni D, Shaya F, Ostersetzer-Biran O (2009) AtnMat2, a nuclear-encoded maturase required for splicing of group-II introns in Arabidopsis mitochondria. *RNA* **15:** 2299-2311

Khadaroo B, Teixeira MT, Luciano P, Eckert-Boulet N, Germann SM, Simon MN, Gallina I, Abdallah P, Gilson E, Geli V, Lisby M (2009) The DNA damage response at eroded telomeres and tethering to the nuclear pore complex. *Nat Cell Biol* **11:** 980-987

Khanna KK, Jackson SP (2001) DNA double-strand breaks: signaling, repair and the cancer connection. *Nat Genet* **27:** 247-254

Kharbanda S, Kumar V, Dhar S, Pandey P, Chen C, Majumder P, Yuan ZM, Whang Y, Strauss W, Pandita TK, Weaver D, Kufe D (2000) Regulation of the hTERT telomerase catalytic subunit by the c-Abl tyrosine kinase. *Curr Biol* **10:** 568-575

Kim NW, Piatyszek MA, Prowse KR, Harley CB, West MD, Ho PL, Coviello GM, Wright WE, Weinrich SL, Shay JW (1994) Specific association of human telomerase activity with immortal cells and cancer. *Science* **266:** 2011-2015

Kim SH, Beausejour C, Davalos AR, Kaminker P, Heo SJ, Campisi J (2004) TIN2 mediates functions of TRF2 at human telomeres. *J Biol Chem* **279:** 43799-43804

Kim SI, Veena, Gelvin SB (2007) Genome-wide analysis of *Agrobacterium* T-DNA integration sites in the *Arabidopsis* genome generated under non-selective conditions. *Plant J* **51:** 779-791

Kinner A, Wu W, Staudt C, Iliakis G (2008) Gamma-H2AX in recognition and signaling of DNA double-strand breaks in the context of chromatin. *Nucleic Acids Res* **36:** 5678-5694

Kipling D (1995) Telomerase: immortality enzyme or oncogene? *Nat Genet* **9:** 104-106

Klobutcher LA, Swanton MT, Donini P, Prescott DM (1981) All gene-sized DNA molecules in four species of hypotrichs have the same terminal sequence and an unusual 3' terminus. *Proc Natl Acad Sci U S A* **78:** 3015-3019

Konig P, Giraldo R, Chapman L, Rhodes D (1996) The crystal structure of the DNA-binding domain of yeast RAP1 in complex with telomeric DNA. *Cell* **85:** 125-136

Koornneef M (2012) My favourite flowering image. *J Exp Bot*

Kramer KM, Haber JE (1993) New telomeres in yeast are initiated with a highly selected subset of TG1-3 repeats. *Genes Dev* **7:** 2345-2356

Krapp S, Kelly G, Reischl J, Weinzierl RO, Matthews S (1998) Eukaryotic RNA polymerase subunit RPB8 is a new relative of the OB family. *Nat Struct Biol* **5:** 110-114

Kumekawa N, Hosouchi T, Tsuruoka H, Kotani H (2001) The size and sequence organization of the centromeric region of *Arabidopsis thaliana* chromosome 4. *DNA Res* **8:** 285-290

Lafarge S, Montane MH (2003) Characterization of Arabidopsis thaliana ortholog of the human breast cancer susceptibility gene 1: AtBRCA1, strongly induced by gamma rays. *Nucleic Acids Res* **31:** 1148-1155

Lamarche BJ, Orazio NI, Weitzman MD (2010) The MRN complex in double-strand break repair and telomere maintenance. *FEBS Lett* **584:** 3682-3695

Lambowitz AM, Zimmerly S (2004) Mobile group II introns. *Annu Rev Genet* **38:** 1-35

Larrivee M, LeBel C, Wellinger RJ (2004) The generation of proper constitutive G-tails on yeast telomeres is dependent on the MRX complex. *Genes Dev* **18:** 1391-1396

Latrick CM, Cech TR (2010) POT1-TPP1 enhances telomerase processivity by slowing primer dissociation and aiding translocation. *EMBO J* **29:** 924-933

Le QH, Wright S, Yu Z, Bureau T (2000) Transposon diversity in Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **97:** 7376-7381

Lee SE, Myung K (2009) Faithful after break-up: suppression of chromosomal translocations. *Cell Mol Life Sci* **66:** 3149-3160

Lei M, Podell ER, Baumann P, Cech TR (2003) DNA self-recognition in the structure of Pot1 bound to telomeric single-stranded DNA. *Nature* **426:** 198-203

Lei M, Podell ER, Cech TR (2004) Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. *Nat Struct Mol Biol* **11:** 1223-1229

Lendvay TS, Morris DK, Sah J, Balasubramanian B, Lundblad V (1996) Senescence mutants of Saccharomyces cerevisiae with a defect in telomere replication identify three additional EST genes. *Genetics* **144:** 1399-1412

Leonardi J, Box JA, Bunch JT, Baumann P (2008) TER1, the RNA subunit of fission yeast telomerase. *Nat Struct Mol Biol* **15:** 26-33

Levy DL, Blackburn EH (2004) Counting of Rif1p and Rif2p on Saccharomyces cerevisiae telomeres regulates telomere length. *Mol Cell Biol* **24:** 10857-10867

Li J, Vaidya M, White C, Vainstein A, Citovsky V, Tzfira T (2005) Involvement of KU80 in T-DNA integration in plant cells. *Proc Natl Acad Sci U S A* **102:** 19231-19236

Lieber MR, Ma Y, Pannicke U, Schwarz K (2003) Mechanism and regulation of human non-homologous DNA end-joining. *Nat Rev Mol Cell Biol* **4:** 712-720

Lin JJ, Zakian VA (1995) An in vitro assay for Saccharomyces telomerase requires EST1. *Cell* **81:** 1127-1135

Lin JJ, Zakian VA (1996) The Saccharomyces CDC13 protein is a single-strand TG1-3 telomeric DNA-binding protein in vitro that affects telomere behavior in vivo. *Proc Natl Acad Sci U S A* **93:** 13760-13765

Linger BR, Price CM (2009) Conservation of telomere protein complexes: shuffling through evolution. *Crit Rev Biochem Mol Biol* **44:** 434-446

Lingner J, Cooper JP, Cech TR (1995) Telomerase and DNA end replication: no longer a lagging strand problem? *Science* **269:** 1533-1534

Lingner J, Hughes TR, Shevchenko A, Mann M, Lundblad V, Cech TR (1997) Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science* **276:** 561-567

Liu D, Safari A, O'Connor MS, Chan DW, Laegeler A, Qin J, Songyang Z (2004) PTOP interacts with POT1 and regulates its localization to telomeres. *Nat Cell Biol* **6:** 673-680

Liu YG, Mitsukawa N, Oosumi T, Whittier RF (1995) Efficient isolation and mapping of Arabidopsis thaliana T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J* **8:** 457-463

Loayza D, De Lange T (2003) POT1 as a terminal transducer of TRF1 telomere length control. *Nature* **423:** 1013-1018

Loayza D, Parsons H, Donigian J, Hoke K, de Lange T (2004) DNA binding features of human POT1: a nonamer 5'-TAGGGTTAG-3' minimal binding site, sequence specificity, and internal binding to multimeric sites. *J Biol Chem* **279:** 13241-13248

Longtine MS, Wilson NM, Petracek ME, Berman J (1989) A yeast telomere binding activity binds to two related telomere sequence motifs and is indistinguishable from RAP1. *Curr Genet* **16:** 225-239

Lopez CR, Ribes-Zamora A, Indiviglio SM, Williams CL, Haricharan S, Bertuch AA (2011) Ku must load directly onto the chromosome end in order to mediate its telomeric functions. *PLoS Genet* **7:** e1002233

Lue NF (2010) Plasticity of telomere maintenance mechanisms in yeast. *Trends Biochem Sci* **35:** 8-17

Lundblad V, Szostak JW (1989) A mutant with a defect in telomere elongation leads to senescence in yeast. *Cell* **57:** 633-643

Lustig AJ, Kurtz S, Shore D (1990) Involvement of the silencer and UAS binding protein RAP1 in regulation of telomere length. *Science* **250:** 549-553

Ly H, Xu L, Rivera MA, Parslow TG, Blackburn EH (2003) A role for a novel 'trans-pseudoknot' RNA-RNA interaction in the functional dimerization of human telomerase. *Genes Dev* **17:** 1078-1083

Lydeard JR, Lipkin-Moore Z, Jain S, Eapen VV, Haber JE (2010) Sgs1 and exo1 redundantly inhibit break-induced replication and de novo telomere addition at broken chromosome ends. *PLoS Genet* **6:** e1000973

Ma Y, Lu H, Tippin B, Goodman MF, Shimazaki N, Koiwai O, Hsieh CL, Schwarz K, Lieber MR (2004) A biochemically defined system for mammalian nonhomologous DNA end joining. *Mol Cell* **16:** 701-713

Makarov VL, Hirose Y, Langmore JP (1997) Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell* **88:** 657-666

Makovets S, Blackburn EH (2009) DNA damage signalling prevents deleterious telomere addition at DNA breaks. *Nat Cell Biol* **11:** 1383-1386

Malek O, Brennicke A, Knoop V (1997) Evolution of trans-splicing plant mitochondrial introns in pre-Permian times. *Proc Natl Acad Sci U S A* **94:** 553-558

Malik HS, Burke WD, Eickbush TH (2000) Putative telomerase catalytic subunits from Giardia lamblia and Caenorhabditis elegans. *Gene* **251:** 101-108

Mangahas JL, Alexander MK, Sandell LL, Zakian VA (2001) Repair of chromosome ends after telomere loss in Saccharomyces. *Mol Biol Cell* **12:** 4078-4089

Marcand S, Brevet V, Mann C, Gilson E (2000) Cell cycle restriction of telomere elongation. *Curr Biol* **10:** 487-490

Marcand S, Gilson E, Shore D (1997) A protein-counting mechanism for telomere length regulation in yeast. *Science* **275:** 986-990

Marrero VA, Symington LS (2010) Extensive DNA end processing by exo1 and sgs1 inhibits break-induced replication. *PLoS Genet* **6:** e1001007

Martin V, Du LL, Rozenzhak S, Russell P (2007) Protection of telomeres by a conserved Stn1-Ten1 complex. *Proc Natl Acad Sci U S A* **104:** 14038-14043

Martina M, Clerici M, Baldo V, Bonetti D, Lucchini G, Longhese MP (2012) A Balance between Tel1 and Rif2 Activities Regulates Nucleolytic Processing and Elongation at Telomeres. *Mol Cell Biol* **32:** 1604-1617

Martinez P, Thanasoula M, Munoz P, Liao C, Tejera A, McNees C, Flores JM, Fernandez-Capetillo O, Tarsounas M, Blasco MA (2009) Increased telomere fragility and fusions resulting from TRF1 deficiency lead to degenerative pathologies and increased cancer in mice. *Genes Dev* **23:** 2060-2075

McClintock B (1938) The Production of Homozygous Deficient Tissues with Mutant Characteristics by Means of the Aberrant Mitotic Behavior of Ring-Shaped Chromosomes. *Genetics* **23:** 315-376

McClintock B (1941) The Stability of Broken Ends of Chromosomes in Zea Mays. *Genetics* **26:** 234-282

McKnight TD, Riha K, Shippen DE (2002) Telomeres, telomerase, and stability of the plant genome. *Plant Mol Biol* **48:** 331-337

McNees CJ, Tejera AM, Martinez P, Murga M, Mulero F, Fernandez-Capetillo O, Blasco MA (2010) ATR suppresses telomere fragility and recombination but is dispensable for elongation of short telomeres by telomerase. *J Cell Biol* **188:** 639-652

McVey M, Lee SE (2008) MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet* **24:** 529-538

Miao ZH, Lam E (1995) Targeted disruption of the TGA3 locus in Arabidopsis thaliana. *Plant J* **7:** 359-365

Miller KM, Rog O, Cooper JP (2006) Semi-conservative DNA replication through telomeres requires Taz1. *Nature* **440:** 824-828

Mirouze M, Paszkowski J (2011) Epigenetic contribution to stress adaptation in plants. *Curr Opin Plant Biol* **14:** 267-274

Mittelsten Scheid O, Afsar K, Paszkowski J (2003) Formation of stable epialleles and their paramutation-like interaction in tetraploid Arabidopsis thaliana. *Nat Genet* **34:** 450-454

Mitton-Fry RM, Anderson EM, Hughes TR, Lundblad V, Wuttke DS (2002) Conserved structure for single-stranded telomeric DNA recognition. *Science* **296:** 145-147

Miyake Y, Nakamura M, Nabetani A, Shimamura S, Tamura M, Yonehara S, Saito M, Ishikawa F (2009) RPA-like mammalian Ctc1-Stn1-Ten1 complex binds to single-stranded DNA and protects telomeres independently of the Pot1 pathway. *Mol Cell* **36:** 193-206

Miyoshi T, Kanoh J, Saito M, Ishikawa F (2008) Fission yeast Pot1-Tpp1 protects telomeres and regulates telomere length. *Science* **320:** 1341-1344

Moore JK, Haber JE (1996) Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in Saccharomyces cerevisiae. *Mol Cell Biol* **16:** 2164-2173

Morin GB (1989) The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. *Cell* **59:** 521-529

Morris GE (2008) The Cajal body. *Biochim Biophys Acta* **1783:** 2108-2115

Moser BA, Chang YT, Kosti J, Nakamura TM (2011) Tel1ATM and Rad3ATR kinases promote Ccq1-Est1 interaction to maintain telomeres in fission yeast. *Nat Struct Mol Biol* **18:** 1408-1413

Moser BA, Subramanian L, Khair L, Chang YT, Nakamura TM (2009) Fission yeast Tel1(ATM) and Rad3(ATR) promote telomere protection and telomerase recruitment. *PLoS Genet* **5:** e1000622

Moshous D, Callebaut I, de Chasseval R, Corneo B, Cavazzana-Calvo M, Le Deist F, Tezcan I, Sanal O, Bertrand Y, Philippe N, Fischer A, de Villartay JP (2001) Artemis, a novel DNA double-strand break repair/V(D)J recombination protein, is mutated in human severe combined immune deficiency. *Cell* **105:** 177-186

Murnane JP (2010) Telomere loss as a mechanism for chromosome instability in human cancer. *Cancer Res* **70:** 4255-4259

Murti KG, Prescott DM (1999) Telomeres of polytene chromosomes in a ciliated protozoan terminate in duplex DNA loops. *Proc Natl Acad Sci U S A* **96:** 14436-14439

Murzin AG (1993) OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J* **12:** 861-867

Nakaoka H, Nishiyama A, Saito M, Ishikawa F (2012) Xenopus laevis Ctc1-Stn1-Ten1 (xCST) protein complex is involved in priming DNA synthesis on single-stranded DNA template in Xenopus egg extract. *J Biol Chem* **287:** 619-627

Negrini S, Ribaud V, Bianchi A, Shore D (2007) DNA breaks are masked by multiple Rap1 binding in yeast: implications for telomere capping and telomerase regulation. *Genes Dev* **21:** 292-302

Newman DR, Kuhn JF, Shanab GM, Maxwell ES (2000) Box C/D snoRNA-associated proteins: two pairs of evolutionarily ancient proteins and possible links to replication and transcription. *RNA* **6:** 861-879

Nigg EA (2001) Mitotic kinases as regulators of cell division and its checkpoints. *Nat Rev Mol Cell Biol* **2:** 21-32

Nosek J, Kosa P, Tomaska L (2006) On the origin of telomeres: a glimpse at the pre-telomerase world. *Bioessays* **28:** 182-190

Nugent CI, Bosco G, Ross LO, Evans SK, Salinger AP, Moore JK, Haber JE, Lundblad V (1998) Telomere maintenance is dependent on activities required for end repair of double-strand breaks. *Curr Biol* **8:** 657-660

Nugent CI, Hughes TR, Lue NF, Lundblad V (1996) Cdc13p: a single-strand telomeric DNA-binding protein with a dual role in yeast telomere maintenance. *Science* **274:** 249-252

Oguchi K, Liu H, Tamura K, Takahashi H (1999) Molecular cloning and characterization of AtTERT, a telomerase reverse transcriptase homolog in Arabidopsis thaliana. *FEBS Lett* **457:** 465-469

Okabe J, Eguchi A, Masago A, Hayakawa T, Nakanishi M (2000) TRF1 is a critical trans-acting factor required for de novo telomere formation in human cells. *Hum Mol Genet* **9:** 2639-2650

Okada-Katsuhata Y, Yamashita A, Kutsuzawa K, Izumi N, Hirahara F, Ohno S (2012) N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. *Nucleic Acids Res* **40:** 1251-1266

Olovnikov AM (1971) [Principle of marginotomy in template synthesis of polynucleotides]. *Dokl Akad Nauk SSSR* **201:** 1496-1499

Olovnikov AM (1973) A theory of marginotomy. The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. *J Theor Biol* **41:** 181-190

O'Malley RC, Ecker JR (2010) Linking genotype to phenotype using the Arabidopsis unimutant collection. *Plant J* **61:** 928-940

O'Sullivan RJ, Karlseder J (2010) Telomeres: protecting chromosomes against genome instability. *Nat Rev Mol Cell Biol* **11:** 171-181

Palm W, de Lange T (2008) How shelterin protects mammalian telomeres. *Annu Rev Genet* **42:** 301-334

Palm W, Hockemeyer D, Kibe T, de Lange T (2009) Functional dissection of human and mouse POT1 proteins. *Mol Cell Biol* **29:** 471-482

Pennaneach V, Kolodner RD (2004) Recombination and the Tel1 and Mec1 checkpoints differentially effect genome rearrangements driven by telomere dysfunction in yeast. *Nat Genet* **36:** 612-617

Pennaneach V, Putnam CD, Kolodner RD (2006) Chromosome healing by de novo telomere addition in *Saccharomyces cerevisiae*. *Mol Microbiol* **59:** 1357-1368

Perroud PF, Quatrano RS (2008) BRICK1 is required for apical cell growth in filaments of the moss Physcomitrella patens but not for gametophore morphology. *Plant Cell* **20:** 411-422

Peterson SE, Stellwagen AE, Diede SJ, Singer MS, Haimberger ZW, Johnson CO, Tzoneva M, Gottschling DE (2001) The function of a stem-loop in telomerase RNA is linked to the DNA repair protein Ku. *Nat Genet* **27:** 64-67

Pickett HA, Cesare AJ, Johnston RL, Neumann AA, Reddel RR (2009) Control of telomere length by a trimming mechanism that involves generation of t-circles. *EMBO J* **28:** 799-809

Polvi A, Linnankivi T, Kivela T, Herva R, Keating JP, Makitie O, Pareyson D, Vainionpaa L, Lahtinen J, Hovatta I, Pihko H, Lehesjoki AE (2012) Mutations in CTC1, encoding the CTS telomere maintenance complex component 1, cause cerebroretinal microangiopathy with calcifications and cysts. *Am J Hum Genet* **90:** 540-549

Porter SE, Greenwell PW, Ritchie KB, Petes TD (1996) The DNA-binding protein Hdf1p (a putative Ku homologue) is required for maintaining normal telomere length in Saccharomyces cerevisiae. *Nucleic Acids Res* **24:** 582-585

Price CM, Boltz KA, Chaiken MF, Stewart JA, Beilstein MA, Shippen DE (2010) Evolution of CST function in telomere maintenance. *Cell Cycle* **9:** 3157-3165

Price CM, Cech TR (1987) Telomeric DNA-protein interactions of Oxytricha macronuclear DNA. *Genes Dev* **1:** 783-793

Prowse KR, Greider CW (1995) Developmental and tissue-specific regulation of mouse telomerase and telomere length. *Proc Natl Acad Sci U S A* **92:** 4818-4822

Przybilski R, Graf S, Lescoute A, Nellen W, Westhof E, Steger G, Hammann C (2005) Functional hammerhead ribozymes naturally encoded in the genome of Arabidopsis thaliana. *Plant Cell* **17:** 1877-1885

Puglisi A, Bianchi A, Lemmens L, Damay P, Shore D (2008) Distinct roles for yeast Stn1 in telomere capping and telomerase inhibition. *EMBO J* **27:** 2328-2339

Qi H, Zakian VA (2000) The Saccharomyces telomere-binding protein Cdc13p interacts with both the catalytic subunit of DNA polymerase alpha and the telomerase-associated est1 protein. *Genes Dev* **14:** 1777-1788

Qiao F, Cech TR (2008) Triple-helix structure in telomerase RNA contributes to catalysis. *Nat Struct Mol Biol* **15:** 634-640

Quatrano RS, McDaniel SF, Khandelwal A, Perroud PF, Cove DJ (2007) Physcomitrella patens: mosses enter the genomic age. *Curr Opin Plant Biol* **10:** 182-189

Raghunathan S, Ricard CS, Lohman TM, Waksman G (1997) Crystal structure of the homo-tetrameric DNA binding domain of Escherichia coli single-stranded DNA-binding protein determined by multiwavelength x-ray diffraction on the selenomethionyl protein at 2.9-A resolution. *Proc Natl Acad Sci U S A* **94:** 6652-6657

Raices M, Verdun RE, Compton SA, Haggblom CI, Griffith JD, Dillin A, Karlseder J (2008) C. elegans telomeres contain G-strand and C-strand overhangs that are bound by distinct proteins. *Cell* **132:** 745-757

Ramakers C, Ruijter JM, Deprez RH, Moorman AF (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci Lett* **339:** 62-66

Ravi M, Chan SW (2010) Haploid plants produced by centromere-mediated genome elimination. *Nature* **464:** 615-618

Ray A, Runge KW (1999) The yeast telomere length counting machinery is sensitive to sequences at the telomere-nontelomere junction. *Mol Cell Biol* **19:** 31-45

Ray S, Karamysheva Z, Wang L, Shippen DE, Price CM (2002) Interactions between telomerase and primase physically link the telomere and chromosome replication machinery. *Mol Cell Biol* **22:** 5859-5868

Reichenbach P, Hoss M, Azzalin CM, Nabholz M, Bucher P, Lingner J (2003) A human homolog of yeast Est1 associates with telomerase and uncaps chromosome ends when overexpressed. *Curr Biol* **13:** 568-574

Rensing SA, Fritzowsky D, Lang D, Reski R (2005) Protein encoding genes in an ancient plant: analysis of codon usage, retained genes and splice sites in a moss, Physcomitrella patens. *BMC Genomics* **6:** 43

Rensing SA, Lang D, Zimmer AD, Terry A, Salamov A, Shapiro H, Nishiyama T, Perroud PF, Lindquist EA, Kamisugi Y, Tanahashi T, Sakakibara K, Fujita T, Oishi K, Shin IT, Kuroki Y, Toyoda A, Suzuki Y, Hashimoto S, Yamaguchi K, Sugano S, Kohara Y, Fujiyama A, Anterola A, Aoki S, Ashton N, Barbazuk WB, Barker E, Bennetzen JL, Blankenship R, Cho SH, Dutcher SK, Estelle M, Fawcett JA, Gundlach H, Hanada K, Heyl A, Hicks KA, Hughes J, Lohr M, Mayer K, Melkozernov A, Murata T, Nelson DR, Pils B, Prigge M, Reiss B, Renner T, Rombauts S, Rushton PJ, Sanderfoot A, Schween G, Shiu SH, Stueber K, Theodoulou FL, Tu H, Van de Peer Y, Verrier PJ, Waters E, Wood A, Yang L, Cove D, Cuming AC, Hasebe M, Lucas S, Mishler BD, Reski R, Grigoriev IV, Quatrano RS, Boore JL (2008) The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. *Science* **319:** 64-69

Ribes-Zamora A, Mihalek I, Lichtarge O, Bertuch AA (2007) Distinct faces of the Ku heterodimer mediate DNA repair and telomeric functions. *Nat Struct Mol Biol* **14:** 301-307

Riehs N, Akimcheva S, Puizina J, Bulankova P, Idol RA, Siroky J, Schleiffer A, Schweizer D, Shippen DE, Riha K (2008) Arabidopsis SMG7 protein is required for exit from meiosis. *J Cell Sci* **121:** 2208-2216

Riehs-Kearnan N, Gloggnitzer J, Dekrout B, Jonak C, Riha K (2012) Aberrant growth and lethality of Arabidopsis deficient in nonsense-mediated RNA decay factors is caused by autoimmune-like response. *Nucleic Acids Res* doi:10.1093/nar/gks195

Riha K, Heacock ML, Shippen DE (2006) The role of the nonhomologous end-joining DNA double-strand break repair pathway in telomere biology. *Annu Rev Genet* **40:** 237-277

Riha K, McKnight TD, Fajkus J, Vyskot B, Shippen DE (2000) Analysis of the G-overhang structures on plant telomeres: evidence for two distinct telomere architectures. *Plant J* **23:** 633-641

Riha K, McKnight TD, Griffing LR, Shippen DE (2001) Living with genome instability: plant responses to telomere dysfunction. *Science* **291:** 1797-1800

Riha K, Shippen DE (2003) Ku is required for telomeric C-rich strand maintenance but not for end-to-end chromosome fusions in Arabidopsis. *Proc Natl Acad Sci U S A* **100:** 611-615

Riha K, Watson JM, Parkey J, Shippen DE (2002) Telomere length deregulation and enhanced sensitivity to genotoxic stress in *Arabidopsis* mutants deficient in Ku70. *EMBO J* **21:** 2819-2826

Rossignol P, Collier S, Bush M, Shaw P, Doonan JH (2007) Arabidopsis POT1A interacts with TERT-V(I8), an N-terminal splicing variant of telomerase. *J Cell Sci* **120:** 3678-3687

Rottbauer W, Saurin AJ, Lickert H, Shen X, Burns CG, Wo ZG, Kemler R, Kingston R, Wu C, Fishman M (2002) Reptin and pontin antagonistically regulate heart growth in zebrafish embryos. *Cell* **111:** 661-672

Ruijter JM, Ramakers C, Hoogaars WM, Karlen Y, Bakker O, van den Hoff MJ, Moorman AF (2009) Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res* **37:** e45

Runge KW, Zakian VA (1990) Properties of the transcriptional enhancer in Saccharomyces cerevisiae telomeres. *Nucleic Acids Res* **18:** 1783-1787

Sabourin M, Tuzon CT, Zakian VA (2007) Telomerase and Tel1p preferentially associate with short telomeres in S. cerevisiae. *Mol Cell* **27:** 550-561

Sabourin M, Zakian VA (2008) ATM-like kinases and regulation of telomerase: lessons from yeast and mammals. *Trends Cell Biol* **18:** 337-346

Salomon S, Puchta H (1998) Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. *EMBO J* **17:** 6086-6095

Samach A, Melamed-Bessudo C, Avivi-Ragolski N, Pietrokovski S, Levy AA (2011) Identification of plant RAD52 homologs and characterization of the Arabidopsis thaliana RAD52-like genes. *Plant Cell* **23:** 4266-4279

Santos JH, Meyer JN, Van Houten B (2006) Mitochondrial localization of telomerase as a determinant for hydrogen peroxide-induced mitochondrial DNA damage and apoptosis. *Hum Mol Genet* **15:** 1757-1768

Santos JL, Alfaro D, Sanchez-Moran E, Armstrong SJ, Franklin FC, Jones GH (2003) Partial diploidization of meiosis in autotetraploid *Arabidopsis thaliana*. *Genetics* **165:** 1533-1540

Sarthy J, Bae NS, Scrafford J, Baumann P (2009) Human RAP1 inhibits non-homologous end joining at telomeres. *EMBO J* **28:** 3390-3399

Schaefer DG (2001) Gene targeting in Physcomitrella patens. *Curr Opin Plant Biol* **4:** 143-150

Schaefer DG (2002) A new moss genetics: targeted mutagenesis in Physcomitrella patens. *Annu Rev Plant Biol* **53:** 477-501

Schaefer DG, Delacote F, Charlot F, Vrielynck N, Guyon-Debast A, Le Guin S, Neuhaus JM, Doutriaux MP, Nogue F (2010) RAD51 loss of function abolishes gene targeting and de-represses illegitimate integration in the moss Physcomitrella patens. *DNA Repair (Amst)* **9:** 526-533

Schnapp G, Rodi HP, Rettig WJ, Schnapp A, Damm K (1998) One-step affinity purification protocol for human telomerase. *Nucleic Acids Res* **26:** 3311-3313

Schneeberger K, Ossowski S, Ott F, Klein JD, Wang X, Lanz C, Smith LM, Cao J, Fitz J, Warthmann N, Henz SR, Huson DH, Weigel D (2011) Reference-guided assembly of four diverse Arabidopsis thaliana genomes. *Proc Natl Acad Sci U S A* **108:** 10249-10254

Schober H, Ferreira H, Kalck V, Gehlen LR, Gasser SM (2009) Yeast telomerase and the SUN domain protein Mps3 anchor telomeres and repress subtelomeric recombination. *Genes Dev* **23:** 928-938

Schoeftner S, Blanco R, Lopez de Silanes I, Munoz P, Gomez-Lopez G, Flores JM, Blasco MA (2009) Telomere shortening relaxes X chromosome inactivation and forces global transcriptome alterations. *Proc Natl Acad Sci U S A* **106:** 19393-19398

Schulz VP, Zakian VA (1994) The saccharomyces PIF1 DNA helicase inhibits telomere elongation and de novo telomere formation. *Cell* **76:** 145-155

Sessions A, Burke E, Presting G, Aux G, McElver J, Patton D, Dietrich B, Ho P, Bacwaden J, Ko C, Clarke JD, Cotton D, Bullis D, Snell J, Miguel T, Hutchison D, Kimmerly B, Mitzel T, Katagiri F, Glazebrook J, Law M, Goff SA (2002) A high-throughput Arabidopsis reverse genetics system. *Plant Cell* **14:** 2985-2994

Seto AG, Livengood AJ, Tzfati Y, Blackburn EH, Cech TR (2002) A bulged stem tethers Est1p to telomerase RNA in budding yeast. *Genes Dev* **16:** 2800-2812

Sfeir A, Kosiyatrakul ST, Hockemeyer D, MacRae SL, Karlseder J, Schildkraut CL, de Lange T (2009) Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication. *Cell* **138:** 90-103

Shakirov EV, Perroud PF, Nelson AD, Cannell ME, Quatrano RS, Shippen DE (2010) Protection of Telomeres 1 is required for telomere integrity in the moss Physcomitrella patens. *Plant Cell* **22:** 1838-1848

Shakirov EV, Shippen DE (2004) Length regulation and dynamics of individual telomere tracts in wild-type Arabidopsis. *Plant Cell* **16:** 1959-1967

Shakirov EV, Surovtseva YV, Osbun N, Shippen DE (2005) The Arabidopsis Pot1 and Pot2 proteins function in telomere length homeostasis and chromosome end protection. *Mol Cell Biol* **25:** 7725-7733

Shay JW, Wright WE (2010) Telomeres and telomerase in normal and cancer stem cells. *FEBS Lett* **584:** 3819-3825

Shiloh Y (2003) ATM and related protein kinases: safeguarding genome integrity. *Nat Rev Cancer* **3:** 155-168

Shippen-Lentz D, Blackburn EH (1990) Functional evidence for an RNA template in telomerase. *Science* **247:** 546-552

Shore D, Bianchi A (2009) Telomere length regulation: coupling DNA end processing to feedback regulation of telomerase. *EMBO J* **28:** 2309-2322

Singer MS, Gottschling DE (1994) TLC1: template RNA component of Saccharomyces cerevisiae telomerase. *Science* **266:** 404-409

Slotkin RK, Martienssen R (2007) Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* **8:** 272-285

Smogorzewska A, van Steensel B, Bianchi A, Oelmann S, Schaefer MR, Schnapp G, de Lange T (2000) Control of human telomere length by TRF1 and TRF2. *Mol Cell Biol* **20:** 1659-1668

Somerville C, Koornneef M (2002) A fortunate choice: the history of Arabidopsis as a model plant. *Nat Rev Genet* **3:** 883-889

Song X, Leehy K, Warrington RT, Lamb JC, Surovtseva YV, Shippen DE (2008) STN1 protects chromosome ends in Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **105:** 19815-19820

Sprung CN, Reynolds GE, Jasin M, Murnane JP (1999) Chromosome healing in mouse embryonic stem cells. *Proc Natl Acad Sci U S A* **96:** 6781-6786

Stellwagen AE, Haimberger ZW, Veatch JR, Gottschling DE (2003) Ku interacts with telomerase RNA to promote telomere addition at native and broken chromosome ends. *Genes Dev* **17:** 2384-2395

Stracker TH, Petrini JH (2011) The MRE11 complex: starting from the ends. *Nat Rev Mol Cell Biol* **12:** 90-103

Studer A, Zhao Q, Ross-Ibarra J, Doebley J (2011) Identification of a functional transposon insertion in the maize domestication gene tb1. *Nat Genet* **43:** 1160-1163

Subramanian L, Moser BA, Nakamura TM (2008) Recombination-based telomere maintenance is dependent on Tel1-MRN and Rap1 and inhibited by telomerase, Taz1, and Ku in fission yeast. *Mol Cell Biol* **28:** 1443-1455

Sun J, Yu EY, Yang Y, Confer LA, Sun SH, Wan K, Lue NF, Lei M (2009) Stn1-Ten1 is an Rpa2-Rpa3-like complex at telomeres. *Genes Dev* **23:** 2900-2914

Surovtseva YV, Churikov D, Boltz KA, Song X, Lamb JC, Warrington R, Leehy K, Heacock M, Price CM, Shippen DE (2009) Conserved telomere maintenance component 1 interacts with STN1 and maintains chromosome ends in higher eukaryotes. *Mol Cell* **36:** 207-218

Surovtseva YV, Shakirov EV, Vespa L, Osbun N, Song X, Shippen DE (2007) Arabidopsis POT1 associates with the telomerase RNP and is required for telomere maintenance. *EMBO J* **26:** 3653-3661

Symington LS, Gautier J (2011) Double-strand break end resection and repair pathway choice. *Annu Rev Genet* **45:** 247-271

Szilard RK, Jacques PE, Laramee L, Cheng B, Galicia S, Bataille AR, Yeung M, Mendez M, Bergeron M, Robert F, Durocher D (2010) Systematic identification of fragile sites via genome-wide location analysis of gamma-H2AX. *Nat Struct Mol Biol* **17:** 299-305

Szostak JW, Blackburn EH (1982) Cloning yeast telomeres on linear plasmid vectors. *Cell* **29:** 245-255

Taggart AK, Teng SC, Zakian VA (2002) Est1p as a cell cycle-regulated activator of telomere-bound telomerase. *Science* **297:** 1023-1026

Takai KK, Kibe T, Donigian JR, Frescas D, de Lange T (2011) Telomere protection by TPP1/POT1 requires tethering to TIN2. *Mol Cell* **44:** 647-659

Talley JM, DeZwaan DC, Maness LD, Freeman BC, Friedman KL (2011) Stimulation of yeast telomerase activity by the ever shorter telomere 3 (Est3) subunit is dependent on direct interaction with the catalytic protein Est2. *J Biol Chem* **286:** 26431-26439

Tang W, Kannan R, Blanchette M, Baumann P (2012) Telomerase RNA biogenesis involves sequential binding by Sm and Lsm complexes. *Nature* **484:** 260-264

Taylor DJ, Podell ER, Taatjes DJ, Cech TR (2011) Multiple POT1-TPP1 proteins coat and compact long telomeric single-stranded DNA. *J Mol Biol* **410:** 10-17

Teixeira MT, Arneric M, Sperisen P, Lingner J (2004) Telomere length homeostasis is achieved via a switch between telomerase- extendible and -nonextendible states. *Cell* **117:** 323-335

Teixeira MT, Gilson E (2007) La sets the tone for telomerase assembly. *Nat Struct Mol Biol* **14:** 261-262

Theimer CA, Blois CA, Feigon J (2005) Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. *Mol Cell* **17:** 671-682

Theimer CA, Feigon J (2006) Structure and function of telomerase RNA. *Curr Opin Struct Biol* **16:** 307-318

Theobald DL, Mitton-Fry RM, Wuttke DS (2003) Nucleic acid recognition by OB-fold proteins. *Annu Rev Biophys Biomol Struct* **32:** 115-133

Theobald DL, Wuttke DS (2004) Prediction of multiple tandem OB-fold domains in telomere end-binding proteins Pot1 and Cdc13. *Structure* **12:** 1877-1879

Ting NS, Yu Y, Pohorelic B, Lees-Miller SP, Beattie TL (2005) Human Ku70/80 interacts directly with hTR, the RNA component of human telomerase. *Nucleic Acids Res* **33:** 2090-2098

Toussaint O, Remacle J, Dierick JF, Pascal T, Frippiat C, Zdanov S, Magalhaes JP, Royer V, Chainiaux F (2002) From the Hayflick mosaic to the mosaics of ageing. Role of stress-induced premature senescence in human ageing. *Int J Biochem Cell Biol* **34:** 1415-1429

Tseng SF, Lin JJ, Teng SC (2006) The telomerase-recruitment domain of the telomere binding protein Cdc13 is regulated by Mec1p/Tel1p-dependent phosphorylation. *Nucleic Acids Res* **34:** 6327-6336

Tseng SF, Shen ZJ, Tsai HJ, Lin YH, Teng SC (2009) Rapid Cdc13 turnover and telomere length homeostasis are controlled by Cdk1-mediated phosphorylation of Cdc13. *Nucleic Acids Res* **37:** 3602-3611

Tuzon CT, Wu Y, Chan A, Zakian VA (2011) The Saccharomyces cerevisiae telomerase subunit Est3 binds telomeres in a cell cycle- and Est1-dependent manner and interacts directly with Est1 in vitro. *PLoS Genet* **7:** e1002060

Tzfati Y, Knight Z, Roy J, Blackburn EH (2003) A novel pseudoknot element is essential for the action of a yeast telomerase. *Genes Dev* **17:** 1779-1788

Tzfira T, Frankman LR, Vaidya M, Citovsky V (2003) Site-specific integration of Agrobacterium tumefaciens T-DNA via double-stranded intermediates. *Plant Physiol* **133:** 1011-1023

Ursic D, Himmel KL, Gurley KA, Webb F, Culbertson MR (1997) The yeast SEN1 gene is required for the processing of diverse RNA classes. *Nucleic Acids Res* **25:** 4778-4785

van Attikum H, Bundock P, Overmeer RM, Lee LY, Gelvin SB, Hooykaas PJ (2003) The *Arabidopsis* AtLIG4 gene is required for the repair of DNA damage, but not for the integration of *Agrobacterium* T-DNA. *Nucleic Acids Res* **31:** 4247-4255

van Attikum H, Hooykaas PJ (2003) Genetic requirements for the targeted integration of Agrobacterium T-DNA in Saccharomyces cerevisiae. *Nucleic Acids Res* **31:** 826-832

van Steensel B, Smogorzewska A, de Lange T (1998) TRF2 protects human telomeres from end-to-end fusions. *Cell* **92:** 401-413

Veldman T, Etheridge KT, Counter CM (2004) Loss of hPot1 function leads to telomere instability and a cut-like phenotype. *Curr Biol* **14:** 2264-2270

Venteicher AS, Abreu EB, Meng Z, McCann KE, Terns RM, Veenstra TD, Terns MP, Artandi SE (2009) A human telomerase holoenzyme protein required for Cajal body localization and telomere synthesis. *Science* **323:** 644-648

Venteicher AS, Artandi SE (2009) TCAB1: driving telomerase to Cajal bodies. *Cell Cycle* **8:** 1329-1331

Venteicher AS, Meng Z, Mason PJ, Veenstra TD, Artandi SE (2008) Identification of ATPases pontin and reptin as telomerase components essential for holoenzyme assembly. *Cell* **132:** 945-957

Verdun RE, Crabbe L, Haggblom C, Karlseder J (2005) Functional human telomeres are recognized as DNA damage in G2 of the cell cycle. *Mol Cell* **20:** 551-561

Verdun RE, Karlseder J (2007) Replication and protection of telomeres. *Nature* **447:** 924-931

Vespa L, Couvillion M, Spangler E, Shippen DE (2005) ATM and ATR make distinct contributions to chromosome end protection and the maintenance of telomeric DNA in Arabidopsis. *Genes Dev* **19:** 2111-2115

Vespa L, Couvillion M, Spangler E, Shippen DE (2005) ATM and ATR make distinct contributions to chromosome end protection and the maintenance of telomeric DNA in Arabidopsis. *Genes Dev* **19:** 2111-2115

Vespa L, Warrington RT, Mokros P, Siroky J, Shippen DE (2007) ATM regulates the length of individual telomere tracts in Arabidopsis. *Proc Natl Acad Sci U S A* **104:** 18145-18150

Virta-Pearlman V, Morris DK, Lundblad V (1996) Est1 has the properties of a single-stranded telomere end-binding protein. *Genes Dev* **10:** 3094-3104

Vizir IY, Mulligan BJ (1999) Genetics of gamma-irradiation-induced mutations in Arabidopsis thaliana: large chromosomal deletions can be rescued through the fertilization of diploid eggs. *J Hered* **90:** 412-417

Wan M, Qin J, Songyang Z, Liu D (2009) OB fold-containing protein 1 (OBFC1), a human homolog of yeast Stn1, associates with TPP1 and is implicated in telomere length regulation. *J Biol Chem* **284:** 26725-26731

Wang F, Podell ER, Zaug AJ, Yang Y, Baciu P, Cech TR, Lei M (2007) The POT1-TPP1 telomere complex is a telomerase processivity factor. *Nature* **445:** 506-510

Wang X, Baumann P (2008) Chromosome fusions following telomere loss are mediated by single-strand annealing. *Mol Cell* **31:** 463-473

Wang Y, Ghosh G, Hendrickson EA (2009) Ku86 represses lethal telomere deletion events in human somatic cells. *Proc Natl Acad Sci U S A* **106:** 12430-12435

Watson JD (1972) Origin of concatemeric T7 DNA. *Nat New Biol* **239:** 197-201

Watson JM, Riha K (2010) Comparative biology of telomeres: where plants stand. *FEBS Lett* **584:** 3752-3759

Watson JM, Shippen DE (2007) Telomere rapid deletion regulates telomere length in Arabidopsis thaliana. *Mol Cell Biol* **27:** 1706-1715

Webb CJ, Zakian VA (2012) Schizosaccharomyces pombe Ccq1 and TER1 bind the 14-3-3-like domain of Est1, which promotes and stabilizes telomerase-telomere association. *Genes Dev* **26:** 82-91

Weinrich SL, Pruzan R, Ma L, Ouellette M, Tesmer VM, Holt SE, Bodnar AG, Lichtsteiner S, Kim NW, Trager JB, Taylor RD, Carlos R, Andrews WH, Wright WE, Shay JW, Harley CB, Morin GB (1997) Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTRT. *Nat Genet* **17:** 498-502

Wellinger RJ (2010) When the caps fall off: responses to telomere uncapping in yeast. *FEBS Lett* **584:** 3734-3740

Wellinger RJ, Wolf AJ, Zakian VA (1993) Saccharomyces telomeres acquire single-strand TG1-3 tails late in S phase. *Cell* **72:** 51-60

West SC (2003) Molecular views of recombination proteins and their control. *Nat Rev Mol Cell Biol* **4:** 435-445

Wilkie AO, Lamb J, Harris PC, Finney RD, Higgs DR (1990) A truncated human chromosome 16 associated with alpha thalassaemia is stabilized by addition of telomeric repeat (TTAGGG)n. *Nature* **346:** 868-871

Williams TL, Levy DL, Maki-Yonekura S, Yonekura K, Blackburn EH (2010) Characterization of the yeast telomere nucleoprotein core: Rap1 binds independently to each recognition site. *J Biol Chem* **285:** 35814-35824

Wong AC, Ning Y, Flint J, Clark K, Dumanski JP, Ledbetter DH, McDermid HE (1997) Molecular characterization of a 130-kb terminal microdeletion at 22q in a child with mild mental retardation. *Am J Hum Genet* **60:** 113-120

Wong JM, Kusdra L, Collins K (2002) Subnuclear shuttling of human telomerase induced by transformation and DNA damage. *Nat Cell Biol* **4:** 731-736

Wotton D, Shore D (1997) A novel Rap1p-interacting factor, Rif2p, cooperates with Rif1p to regulate telomere length in Saccharomyces cerevisiae. *Genes Dev* **11:** 748-760

Wright WE, Tesmer VM, Huffman KE, Levene SD, Shay JW (1997) Normal human chromosomes have long G-rich telomeric overhangs at one end. *Genes Dev* **11:** 2801-2809

Wu Y, Zakian VA (2011) The telomeric Cdc13 protein interacts directly with the telomerase subunit Est1 to bring it to telomeric DNA ends in vitro. *Proc Natl Acad Sci U S A* **108:** 20362-20369

Xin H, Liu D, Wan M, Safari A, Kim H, Sun W, O'Connor MS, Songyang Z (2007) TPP1 is a homologue of ciliate TEBP-beta and interacts with POT1 to recruit telomerase. *Nature* **445:** 559-562

Xu L, Petreaca RC, Gasparyan HJ, Vu S, Nugent CI (2009) TEN1 is essential for CDC13-mediated telomere capping. *Genetics* **183:** 793-810

Xu Y, Price BD (2011) Chromatin dynamics and the repair of DNA double strand breaks. *Cell Cycle* **10:** 261-267

Yamazaki H, Tarumoto Y, Ishikawa F (2012) Tel1(ATM) and Rad3(ATR) phosphorylate the telomere protein Ccq1 to recruit telomerase and elongate telomeres in fission yeast. *Genes Dev* **26:** 241-246

Yamazaki H, Tarumoto Y, Ishikawa F (2012) Tel1(ATM) and Rad3(ATR) phosphorylate the telomere protein Ccq1 to recruit telomerase and elongate telomeres in fission yeast. *Genes Dev* **26:** 241-246

Yang CP, Chen YB, Meng FL, Zhou JQ (2006) Saccharomyces cerevisiae Est3p dimerizes in vitro and dimerization contributes to efficient telomere replication in vivo. *Nucleic Acids Res* **34:** 407-416

Ye JZ, de Lange T (2004) TIN2 is a tankyrase 1 PARP modulator in the TRF1 telomere length control complex. *Nat Genet* **36:** 618-623

Yen WF, Chico L, Lei M, Lue NF (2011) Telomerase regulatory subunit Est3 in two Candida species physically interacts with the TEN domain of TERT and telomeric DNA. *Proc Natl Acad Sci U S A* **108:** 20370-20375

Yoshiyama K, Conklin PA, Huefner ND, Britt AB (2009) Suppressor of gamma response 1 (SOG1) encodes a putative transcription factor governing multiple responses to DNA damage. *Proc Natl Acad Sci U S A* **106:** 12843-12848

Yu W, Han F, Gao Z, Vega JM, Birchler JA (2007) Construction and behavior of engineered minichromosomes in maize. *Proc Natl Acad Sci U S A* **104:** 8924-8929

Yu W, Lamb JC, Han F, Birchler JA (2006) Telomere-mediated chromosomal truncation in maize. *Proc Natl Acad Sci U S A* **103:** 17331-17336

Zappulla DC, Cech TR (2004) Yeast telomerase RNA: a flexible scaffold for protein subunits. *Proc Natl Acad Sci U S A* **101:** 10024-10029

Zappulla DC, Cech TR (2006) RNA as a flexible scaffold for proteins: yeast telomerase and beyond. *Cold Spring Harb Symp Quant Biol* **71:** 217-224

Zappulla DC, Goodrich K, Cech TR (2005) A miniature yeast telomerase RNA functions in vivo and reconstitutes activity in vitro. *Nat Struct Mol Biol* **12:** 1072-1077

Zappulla DC, Roberts JN, Goodrich KJ, Cech TR, Wuttke DS (2009) Inhibition of yeast telomerase action by the telomeric ssDNA-binding protein, Cdc13p. *Nucleic Acids Res* **37:** 354-367

Zaug AJ, Linger J, Cech TR (1996) Method for determining RNA 3' ends and application to human telomerase RNA. *Nucleic Acids Res* **24:** 532-533

Zellinger B, Akimcheva S, Puizina J, Schirato M, Riha K (2007) Ku suppresses formation of telomeric circles and alternative telomere lengthening in Arabidopsis. *Mol Cell* **27:** 163-169

Zellinger B, Riha K (2007) Composition of plant telomeres. *Biochim Biophys Acta* **1769:** 399-409

Zhang ML, Tong XJ, Fu XH, Zhou BO, Wang J, Liao XH, Li QJ, Shen N, Ding J, Zhou JQ (2010) Yeast telomerase subunit Est1p has guanine quadruplex-promoting activity that is required for telomere elongation. *Nat Struct Mol Biol* **17:** 202-209

Zhang Q, Kim NK, Feigon J (2011) Architecture of human telomerase RNA. *Proc Natl Acad Sci U S A* **108:** 20325-20332

Zhang W, Durocher D (2010) De novo telomere formation is suppressed by the Mec1-dependent inhibition of Cdc13 accumulation at DNA breaks. *Genes Dev* **24:** 502-515

Zhang X, Henriques R, Lin SS, Niu QW, Chua NH (2006) *Agrobacterium*-mediated transformation of Arabidopsis thaliana using the floral dip method. *Nat Protoc* **1:** 641-646

Zhang XL, Hong GF (2000) Preferential Location of MITEs in Rice Genome. *Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai)* **32:** 223-228

Zhong F, Savage SA, Shkreli M, Giri N, Jessop L, Myers T, Chen R, Alter BP, Artandi SE (2011) Disruption of telomerase trafficking by TCAB1 mutation causes dyskeratosis congenita. *Genes Dev* **25:** 11-16

Zhong Z, Shiue L, Kaplan S, de Lange T (1992) A mammalian factor that binds telomeric TTAGGG repeats in vitro. *Mol Cell Biol* **12:** 4834-4843

Zhou J, Monson EK, Teng SC, Schulz VP, Zakian VA (2000) Pif1p helicase, a catalytic inhibitor of telomerase in yeast. *Science* **289:** 771-774

Zimmerly S, Hausner G, Wu X (2001) Phylogenetic relationships among group II intron ORFs. *Nucleic Acids Res* **29:** 1238-1250

Zou L, Elledge SJ (2003) Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science* **300:** 1542-1548

APPENDIX I

DYSKERIN IS A COMPONENT OF THE ARABIDOPSIS TELOMERASE RNP REQUIRED FOR TELOMERE MAINTENANCE*

Summary

Dyskerin binds the H/ACA box of human telomerase RNA and is a core telomerase subunit required for RNP biogenesis and enzyme function *in vivo*. Missense mutations in dyskerin result in dyskeratosis congenita (DC), a complex syndrome characterized by bone marrow failure, telomerase enzyme deficiency and progressive telomere shortening. Here we demonstrate that dyskerin also contributes to telomere maintenance in *Arabidopsis*. We report that both AtNAP57, the *Arabidopsis* dyskerin homolog and AtTERT, the telomerase catalytic subunit, accumulate in the plant nucleolus and AtNAP57 associates with active telomerase RNP particles in an RNA-dependent manner. Furthermore, AtNAP57 interacts *in vitro* with AtPOT1a, a novel component of *Arabidopsis* telomerase. Although a null mutation in *AtNAP57* is lethal, AtNAP57, like AtTERT, is not haploinsufficient for telomere maintenance in *Arabidopsis*. However, introduction of an *AtNAP57* allele containing a T66A mutation decreased telomerase activity *in vitro*, disrupted telomere length regulation on individual chromosome ends *in vivo* and established a new shorter telomere length set point. These

---

*Reprinted with permission from Kannan K., Nelson A.D.L., and Shippen D.E. 2008. Dyskerin Is A Component Of The Arabidopsis Telomerase RNP Required For Telomere Maintenance, *Molecular and Cellular Biology* **28**, 2332-2341.

results imply that T66A NAP57 behaves as a dominant negative inhibitor of telomerase. We conclude that dyskerin is a conserved component of the telomerase RNP complex in higher eukaryotes that is required for maximal enzyme activity *in vivo*.

INTRODUCTION

An essential step in the maturation of ribosomal RNA is the conversion of uridine to pseudouridine by H/ACA ribonucleoproteins (RNPs) (30). Components of H/ACA RNPs include small nucleolar RNAs (snoRNAs), Gar1, Nhp2, Nop10, and the pseudouridine synthase, dyskerin. Dyskerin is an essential gene and its loss results in embryonic lethality in mice (18). In addition to its role in rRNA maturation, dyskerin also binds the H/ACA box of human telomerase RNA (hTR) and is involved in hTR processing and stabilization (6, 32). Mass spectrometry studies indicate that the core telomerase complex is composed of a dimer of a catalytic telomerase reverse transcriptase (TERT), TR (which acts as a template for TERT), and dyskerin (7). Notably, the dyskerin homolog in yeast, Cbf5p, is not stably associated with the telomerase RNA (9), and a different constellation of proteins is required for telomerase RNP biogenesis and enzyme function in this organism (8).

Mutations in human dyskerin are the cause of X-linked dyskeratosis congenita (DC), a rare disease that affects regenerative tissues and is characterized by abnormal skin pigmentation and bone marrow failure (20). Patients suffering from X-linked DC have shorter telomeres relative to age-matched controls (32). Most mutations in patients with X-linked DC cluster around the PUA (pseudouridine synthase and archeosine

transglycosylase) domain of dyskerin, which is responsible for RNA binding (35). One of the most commonly identified dyskerin mutations, A353V, perturbs rRNA pseudouridylation and also results in reduced levels of TR, decreased telomerase activity, and shorter telomeres in mouse embryonic stem cells (33). Similarly, hypomorphic mice that express low levels of dyskerin display the clinical symptoms of DC and exhibit shorter telomeres, but only in later generations (39). While rRNA processing is affected in some dyskerin mutants, the T66A mutation in humans appears to exclusively affect the telomerase-associated functions of dyskerin (32).

Recent data indicate that bone marrow disease can also arise through reduction of other telomerase RNP constituents. Heterozygous mutations in hTR, which reduce its accumulation and perturbs its structure, lead to an autosomal dominant form of DC through haploinsufficiency of the RNA subunit (5, 14, 45). Similarly, haploinsufficiency of TERT has been implicated in DC and in aplastic anemia (1, 47, 51). Limiting abundance of telomerase subunits may help to facilitate the fine balance of telomerase repression and activation associated with differentiated cells and their stem cell progenitors (15).

The flowering plant *Arabidopsis* is a useful model for telomere biology (29). In contrast to mouse, *Arabidopsis* telomere tracts are relatively short (2-5 kb) and are abutted by unique sequences on most chromosome arms (19), making it possible to study the dynamics of individual telomeres. Moreover, *Arabidopsis* is exceptionally tolerant to telomere dysfunction and genome instability. Disruption of *AtTERT* results in a slow, but progressive loss of telomeric DNA (12). Beginning in the sixth generation

(G6), *tert* mutants exhibit a low level of end-to-end chromosome fusions and the onset of growth and developmental defects (37). Remarkably, plants survive for up to five more generations with worsening phenotypes until they ultimately arrest growth in a miniature, de-differentiated state unable to produce a germline (36).

Aside from the presence of AtTERT, little is known about telomerase RNP composition and biogenesis in plants. The TR subunit has not yet been identified in any plant species, owing to the rapid evolution of the TR nucleotide sequence. However, recent studies indicate that AtPOT1a, an OB-fold containing protein whose counterparts in yeast and mammals associate with telomeres (3), functions as a telomerase RNP accessory factor in *Arabidopsis* (43). This observation implies that the composition and/or role of telomerase subunits may vary among higher eukaryotes.

Arabidopsis encodes a dyskerin homolog, *AtNAP57* (25, 28) and here we examine its contribution to telomerase biochemistry and telomere maintenance. We demonstrate that AtNAP57 localizes to the nucleolus along with AtTERT and associates with enzymatically active telomerase RNP particles in an RNA-dependent fashion. Although a null mutation is lethal, AtNAP57 is not haploinsufficient for telomere maintenance. However, transgenic plants carrying an *AtNAP57* allele with a T66A mutation exhibit decreased telomerase activity *in vitro* and *in vivo*, deregulated telomere tracts on individual chromosome ends, and shorter, but stable telomeres. We conclude that dyskerin is a conserved component of the telomerase RNP in multicellular organisms that is required for telomere maintenance.

MATERIALS AND METHODS

Plant materials, genotyping, c-DNA synthesis, site-directed mutagenesis and transformation

Arabidopsis seeds with T-DNA insertions in the *AtNAP57* (SALK_031065) and *AtKU70* (SALK_ 123114) genes were purchased from the *Arabidopsis* Biological Resource Center (Ohio State University, Columbus, Ohio), cold-treated overnight at 4°C, and then placed in an environmental growth chamber and grown under a 16 h light/8 h dark photoperiod at 23°C. *Arabidopsis* suspension culture cells were maintained as described (31). Siliques from wild-type and AtNAP57 heterozygotes were dissected 10 days after fertilization and photographed using a Zeiss Axiocam digital camera coupled to a Zeiss microscope.

For genotyping, DNA was extracted from flowers and PCR was performed with the following sets of primers for *AtNAP57*: D5: 5' GTCGACATCTCACACTCGAA 3' and
D 8: 5' GTCTCACTTTGTTCCAGAGT 3' and for *AtKU70*: Ku 1: 5' TTACTTTGTTGTTTCGGGTGC 3' and Ku 2: 5' CTCTTGGCAAGTACACGCTTC 3'.

Total RNA was extracted from 0.5 g of plant tissue using Tri Reagent solution (Sigma). cDNAs were synthesized from total RNA using Superscript III reverse transcriptase (Invitrogen). Oligo dT primers were incubated with 2 µg of total RNA in the supplied buffer at 65°C for 5 min. Reverse transcription was carried with 100 U of Superscript III at 55°C for 60 min. RNA was degraded with RNase H (USB). For

amplifying AtNAP57, we used primers D5 (above) and D2: 5’ GCCATCACAGATGTTGTCATC 3’.

The genomic copy of AtNAP57 (AtNAP57 cDNA + 1kb promoter) was amplified by PCR and ligated into a binary vector pCBK05 (38) lacking the 35S CaMV promoter. To generate the T66A mutation, site-directed mutagenesis was performed with *Pfu* turbo polymerase (Stratagene) on the genomic version of AtNAP57 in pCBKO5 using the primers M1: 5’ CCTCAACGTCCGTGCCGGTCAC 3’ and M2: 5’ GTGACCGGCACGGACGTTGAGG 3’ according to the manufacturer’s guidelines. The construct was introduced into *Agrobacterium tumefaciens* strain GV3101. Transformation of *AtNAP57* heterozygous plants was performed by the *in planta* method as described in (38). Transformants were selected on 0.5 Murashige and Skoog basal medium supplemented with 20 mg/L of phosphinothricine (Crescent Chemical) and Kanamycin (50 µg/mL) and then genotyped. To generate epitope-tagged protein, AtNAP57 cDNA was amplified and ligated it into pCBKO5 with an N-terminal 3X FLAG tag. This construct was transformed into *Agrobacterium* and then transformed into wild-type plants as described above.

TRF analysis, PETRA, TRAP and Quantitative TRAP assays

DNA from individual whole plants was extracted and TRF analysis was performed with *Tru1*I (Fermentas) restriction enzyme and [^{32}P] 5' end-labeled $(T_3AG_3)_4$ oligonucleotide as a probe (12). The peak value for bulk telomere length (Figure 5D) was determined by ImageQuant software. PETRA analysis was conducted on DNA

from whole plants as described (49). TRAP protein extraction and assays were performed on flowers as previously described (11).

Real-time quantitative TRAP was performed as described (22), but with the following modifications. 10.5 µl of a 4.8 ng/µl protein extract dilution, 1 µl of 10 µM forward primer (5' CACTATCGACTACGCGATCAG 3'), and 12.5 µl SYBR Green PCR Master Mix (NEB) were incubated at 37°C for 45 min. 1 µl of reverse primer (5' CCCTAAACCCTAAACCCTAAA 3') was added and products were amplified for 35 PCR cycles with 30 sec at 95°C and 90 sec at 60°C. Threshold cycle values (C_t) were calculated using an iCycler iQ thermal cycler (BIO-RAD) and the supplied Optical System Software. Samples were analyzed in triplicate, with inactivated samples and lysis buffer serving as negative controls.

Western blotting, immunoprecipitation and immunofluorescence

Plant extracts were made by grinding 0.3 g of flowers in buffer A (50 mM Tris-Cl pH7.5, 10 mM $MgCl_2$, 100 mM NaCl, 1 mM EDTA, 10% glycerol, 1 mM DTT and plant protease inhibitors (SIGMA)). Western blotting was performed with a 1:1000 dilution of anti-FLAG antibody (SIGMA) and a 1:10,000 dilution of HRP-conjugated anti-Mouse IgG (SIGMA). For immunoprecipitation, 50 µl of α-FLAG beads (SIGMA) were washed four times with buffer A and incubated with 500 µl of extract for 2 h at 4°C. Beads were then washed three times with buffer A and eluted using the 3X FLAG peptide for 30 min.

A peptide antibody against AtTERT was raised in rabbits and affinity purified (Covance). The peptide used was N-CIKHKRTLSVHENKRKRDDNVQP

corresponding to residues 180-202 of AtTERT. Peptide antibodies against mouse dyskerin (33) were a gift from Dr. Monica Bessler.

Arabidopsis suspension culture extracts were made as above, diluted in buffer W-100 (20 mM TrisOAc pH 7.5, 10% glycerol, 1 mM EDTA, 5 mM $MgCl_2$, 0.2 M NaCl, 1% NP-40, 0.5 mM sodium deoxycholate and 100 mM potassium glutamate) and pre-cleared with Protein-A agarose (PIERCE). Extracts were incubated with antibody and pre-blocked beads. Beads were washed three times with W-300 (W-100 containing 300 mM potassium glutamate) and once with TMG (10 mM TrisOAc pH 7.5, 1 mM $MgCl_2$ and 10% glycerol). The beads were then used for either TRAP or western blotting. Proteins were expressed in rabbit reticulocyte lysate (RRL) (Promega) according to the manufacturer's protocol and used in immunoprecipitation experiments as above. The fraction of enzymatically active telomerase particles that was associated with AtNAP57 was determined by calculating the efficiency with which the TERT antibody immunoprecipitated telomerase activity in a TRAP assay (relative to input) using ImageQuant software. This value was compared to the amount of AtNAP57 signal obtained by western blot analysis following immunoprecipitation of these same samples using QuantityOne software.

For immunofluorescence, *Arabidopsis* suspension culture cells and floral buds were fixed with 3.5% formaldehyde in 1X PBS for 30 min and then washed with 1X PBS. Cells were soaked in 1X citric buffer (10 mM sodium citrate and 10 mM EDTA) for 10 min. Citric buffer was removed and enzyme mix (1% pectinase, 4% cellulose and 1% macerozyme) was added and incubated at 37°C for 40 min. Cells were rinsed with

1X PBS and spun down onto poly-lysine coated slides in a swinging bucket rotor centrifuge for 3 min at 300 x*g*. Slides were removed from the centrifuge and immediately soaked in 1X PBS with 0.2% Triton X-100 for 30 min to permeabilize the cells. Slides were washed with 1X PBS and treated with Image-IT solution (Molecular probes). Primary antibodies (1:200 dilution for AtTERT and 1:400 dilution for dyskerin) were added to the slides, covered with a plastic coverslip, and incubated overnight at room temperature. After washes in 1X PBS and PI, secondary antibody (Goat anti-rabbit IgG conjugated to Texas Red 1:200 dilution) was added and incubated for 4 h. Slides were washed, Vectashield containing DAPI was applied, and images were captured using a CCD camera coupled to a Zeiss epifluorescent microscope.

Co-immunoprecipiation and yeast two-hybrid assays

Co-immunoprecipitation was performed as previously described (23) using full-length AtTERT, AtNAP57, AtKU70, AtKU80 and AtPOT1a proteins expressed in RRL. All components of the yeast two-hybrid system were obtained from Clontech Laboratories. AtNAP57 was subcloned from FLAG-AtNAP57-pCBKO5 into pAS2-1. KU70-pAS2-1, KU80-pAS2-1 and NAP57-pAS2-1 were transformed into the yeast strain AH109. AtKU80 and AtPOT1a were cloned into the prey vector pACT2 and then transformed into Y187 strain. Yeast mating assays were performed as detailed in the Clontech yeast protocols handbook (# PT3024-1). Double selection (SD/-leucine/-tryptophan) was used to obtain diploids and triple selection (SD/- leucine/-tryptophan/–histidine) was used to screen for interaction. To confirm interactions, β-galactosidase

assays (colony-lift filter assay) were performed on the colonies that grew on triple selection plates and development of the blue color was followed for several hours.

RESULTS

Characterization of Arabidopsis AtNAP57

AtNAP57 is encoded by a single gene (At3g57150) on the third *Arabidopsis* chromosome (28) and has only a single exon and no introns. RT-PCR experiments revealed that expression of the 1.6 kb AtNAP57 transcript is ubiquitous (Figure A1-1A), as for mammalian dyskerin (16). AtNAP57 mRNA translates to a highly basic protein with a molecular mass of 63 kDa. A heterologous antibody directed at mouse dyskerin (33) immunoprecipitated recombinant AtNAP57 expressed in rabbit reticulocyte lysate (RRL) (Figure A1-1B) as well as the endogenous plant protein (see Figure A1-2D). This antibody was used to examine the subcellular localization of AtNAP57 in *Arabidopsis* suspension culture cells. Consistent with a role for the plant AtNAP57 in rRNA processing, a bright signal for AtNAP57 was detected exclusively in the nucleolus (Figure A1-1C). A similar finding was observed by Lermontova et al. in a recent study of *Arabidopsis* dyskerin (25). We next asked whether AtTERT also localized to this compartment using a peptide antibody raised against AtTERT. The anti-TERT antibody recognized recombinant AtTERT expressed in RRL (Figure 1B) as well as the endogenous protein from suspension culture extracts (data not shown).

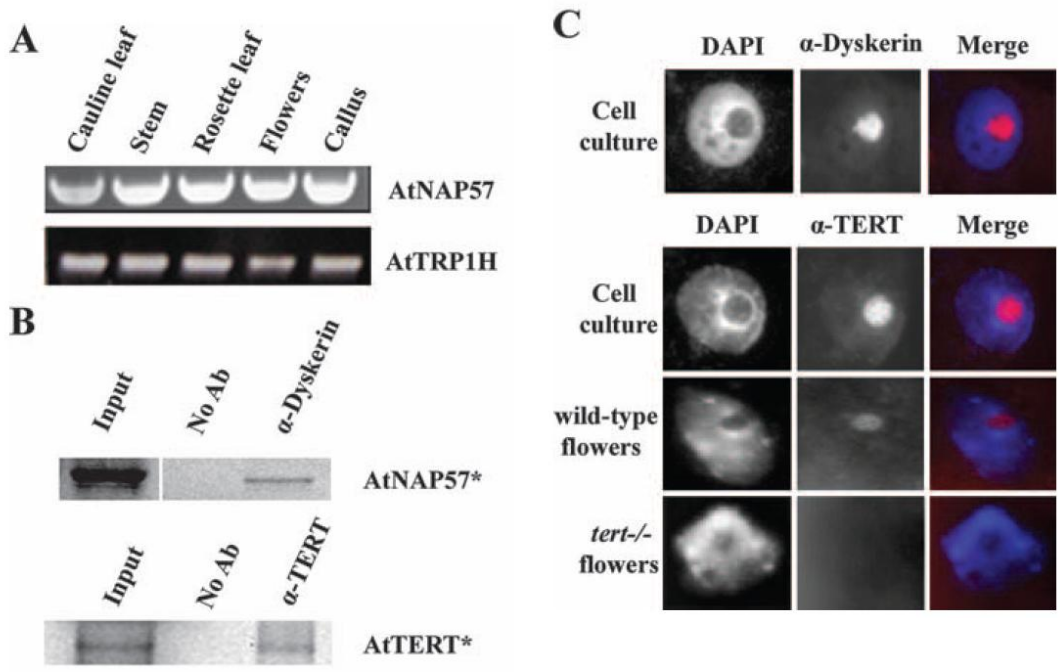


Fig A1-1. Expression and localization of AtNAP57. (**A**) RT-PCR analysis of the AtNAP57 transcript in different plant tissues. AtTRP1H encodes a putative double-strand telomere binding protein (23) and was used as a loading control. (**B**) Recombinant AtNAP57 and AtTERT proteins were expressed in RRL and labeled with [35S]methionine (*). Proteins were immunoprecipitated with an antibody (Ab) raised against mouse dyskerin (-dyskerin) or an antibody raised against an N-terminal peptide in AtTERT (-TERT). Relevant lanes are shown. (**C**) Immunolocalization of AtNAP57 and AtTERT in Arabidopsis suspension culture cells and in floral buds. Nuclei were stained with DAPI or the antibodies discussed above.

Immunolocalization experiments reveal that AtTERT, like AtNAP57, localized to the nucleolus (Figure A1-1C). Nucleolar localization of AtTERT was not detected in *tert*$^{-/-}$ flowers indicating that staining is specific. These findings imply that telomerase biogenesis may occur in the plant nucleolus.

AtNAP57 is a component of the Arabidopsis telomerase RNP.

We asked if AtNAP57 physically associates with the *Arabidopsis* telomerase RNP. A fusion construct was generated containing three copies of the FLAG epitope inserted at the N-terminus of AtNAP57 coding region under the control of the robust cauliflower mosaic virus 35S promoter. This construct was transformed into wild-type *Arabidopsis* and transformants were analyzed by western blotting using a FLAG antibody. Approximately 1 in 10 of the herbicide resistant transformants generated detectable levels of the FLAG-AtNAP57 protein (Figure A1-2A; data not shown). These plants were used for further study. Complexes containing FLAG-AtNAP57 were immunoprecipitated from transgenic plants and eluted using 3X FLAG peptide. As expected, AtNAP57 was immunoprecipitated from transgenic plants, but not from their wild-type counterparts (Figure A1- 2B).

To monitor AtNAP57 association with telomerase, the telomere repeat amplification protocol (TRAP) was performed on FLAG-AtNAP57 and wild-type immunoprecipitates. Telomerase activity was immunoprecipitated from FLAG-AtNAP57 plants, but not from wild-type plants lacking FLAG-AtNAP57 (Figure A1-2C). To verify that the AtNAP57 interaction with telomerase was specific, we performed a reciprocal

immunoprecipitation experiment using the TERT peptide antibody to pull-down AtNAP57 from suspension culture cell extract. Telomerase activity was immunoprecipitated, and as expected, pre-treatment of the extract with RNase A abolished telomerase activity (Figure A1-2D). Notably, western blot analysis revealed a strong enrichment of AtNAP57 in the α-TERT immunoprecipitate, but not when the extract was pre-treated with RNase A prior to immunoprecipitation (Figure A1-2D, bottom panel). To determine the relative amount of telomerase that was associated with AtNAP57, we compared the efficiency of the AtTERT IP, which was determined to be approximately 10% from the experiment in Figure A1-1B, to the amount of AtNAP57 recovered. From these data, we estimate that more than 90% of the active telomerase precipitated by the TERT antibody is associated with AtNAP57. These findings indicate that AtNAP57 is associated with catalytically active telomerase RNP particles and that this interaction requires RNA.

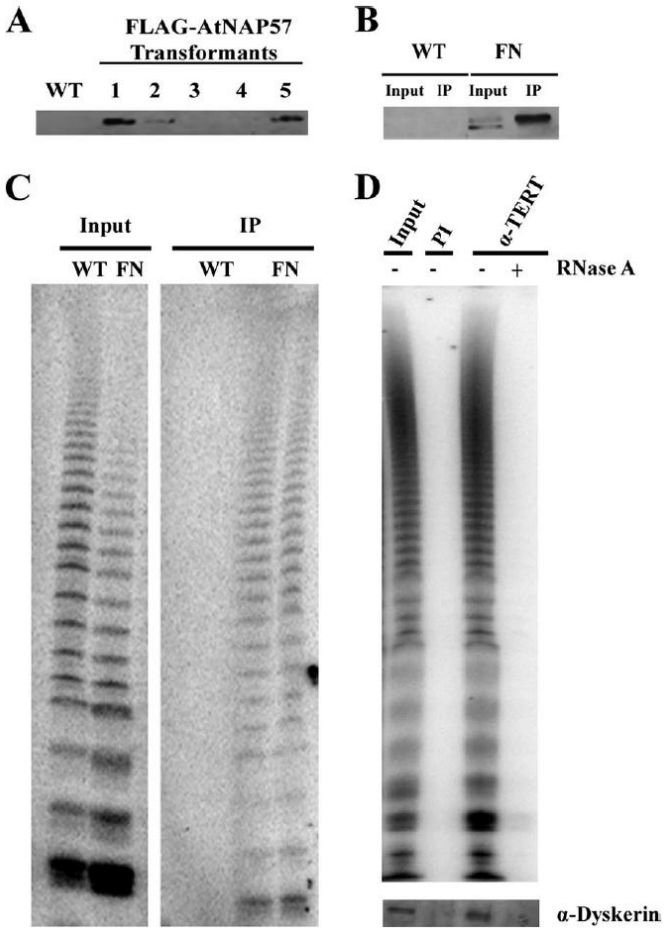


Fig A1-2. AtNAP57 associates with Arabidopsis telomerase RNP. (**A**) Western blot analysis with FLAG antibody on plant extracts from the wild type (WT) or transformants bearing FLAG-tagged AtNAP57 (FN). (**B**) Western blot analysis of input or immunoprecipitates (IP) obtained with FLAG antibody on extracts from the WT and FLAG-tagged AtNAP57 transformants. Four percent of input and 30% of IP was loaded on the gel. (**C**) TRAP assay results for WT and FN extracts before (input) or after immunoprecipitation (IP) with FLAG antibody. Immunoprecipitates were assayed in duplicate. (D) Top panel, TRAP assay results for cell culture extracts immunoprecipitated with preimmune serum (PI) and anti-TERT (-TERT) peptide antibody. The sample shown in the far-right lane was pretreated with 100 g/ml of RNase A prior to immunoprecipitation. Bottom panel, -TERT immunoprecipitates were subjected to Western blot analysis using the dyskerin antibody (-dyskerin). Fifteen percent of input and 60% of IP was loaded on the gel.

In mammals, dyskerin primarily associates with telomerase through TR. Because TR has not yet been identified in *Arabidopsis*, we asked whether AtNAP57 interacts with the known telomerase-associated proteins *in vitro* using co-immunoprecipitation. As expected in our control reactions (38) we detected the formation of AtKU70-AtKU80 heterodimers, while no interaction was observed for AtKU70 alone (Figure A1-3A). We failed to observe binding of T7-tagged AtNAP57 to radiolabeled full-length TERT protein (Figure A1-3A), and similarly in reciprocal co-immunoprecipitation experiments with T7-tagged AtTERT and labeled AtNAP57, no interaction was detected (data not shown). Furthermore, we did not detect binding of AtNAP57 to segments corresponding to the N-terminus, middle and C-terminus of TERT. Although there was a high background in the AtPOT1a control reaction with beads alone, AtPOT1a abundance was reproducibly higher in the immunoprecipitate of T7-tagged AtNAP57(Figure A1-3A; data not shown). The AtNAP57-POT1a interaction appears to be specific as a closely related protein, AtPOT1b (40) was not precipitated with T7-tagged AtNAP57 (data not shown). We confirmed the AtNAP57-AtPOT1a interaction using a yeast two-hybrid mating assay. Yeast strains containing different plasmids were mated and diploids were selected on triple selection (Figure A1-3B). To monitor reporter gene activity, β−galactosidase assays were performed. Blue staining, indicative of interaction, was observed for AtKU70-AtKU80 within one hour, and after four hours staining was detected for AtNAP57-AtPOT1a (Figure A1-3B). Thus, the interaction between AtNAP57 and AtPOT1a is weak, but specific. Taken together, these data argue that AtNAP57 is physically associated with the *Arabidopsis* telomerase RNP.

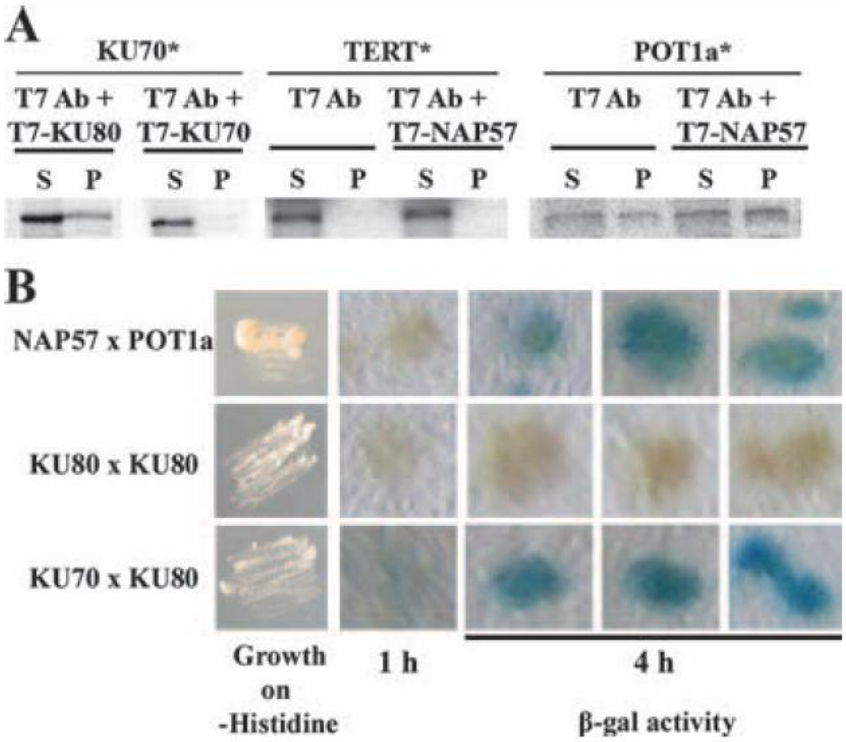


Fig A1-3. FIG. 3. AtNAP57 weakly associates with AtPOT1a. (A) Coimmunoprecipitation experiments were performed with the full-length recombinant AtKU70, AtTERT, and AtPOT1a proteins, labeled using [35S]methionine (*****), and T7-tagged AtKU80, AtKU70, and AtNAP57. Proteins were incubated with either T7 antibody (Ab) beads (control) or T7 beads and the indicated T7-tagged unlabeled proteins. The supernatant (S) and pellet (P) fractions were loaded in equal amounts. (B) Results of yeast two-hybrid analysis are shown. The indicated yeast crosses were performed and plated on medium lacking leucine, tryptophan, and histidine. Results of a colony lift-galactosidase (-gal) assay are shown. The blue color is indicative of protein interaction.

AtNAP57 is not haploinsufficient for telomere maintenance in Arabidopsis.

To investigate the role of AtNAP57 in telomere maintenance, we obtained a mutant line (SALK_031065) carrying a T-DNA insertion in the extreme 5' end of the gene corresponding to the 18th amino acid of the AtNAP57 ORF (Figure A1-4A). Although we genotyped a population of more than 50 progeny from this line, we did not recover any homozygous mutants (data not shown). Dissection of siliques (seed pods) from the heterozygous mutants revealed a reduced seed set in which approximately 25% of the seeds failed to form viable embryos (Figure A1-4B) This outcome implies that AtNAP57 is an essential gene, a conclusion consistent with a recent report for plants with homozygous mutation in AtNAP57 (25).

Heterozygous AtNAP57 mutants were indistinguishable from wild-type in their growth and development over successive plant generations. Furthermore, terminal restriction fragment (TRF) analysis conducted on ten first generation (G1) *nap57*$^{+/-}$ mutants revealed some variability in bulk telomere length, but in all cases, telomeres were in the wild-type 2-5 kb size range (Figure A1-4C, left panel; data not shown). A similar result was obtained when telomeres from G2 *nap57*$^{+/-}$ were examined (Figure 4C, right panel; data not shown). These data imply that AtNAP57 is not haploinsufficient for telomere length maintenance in *Arabidopsis.*

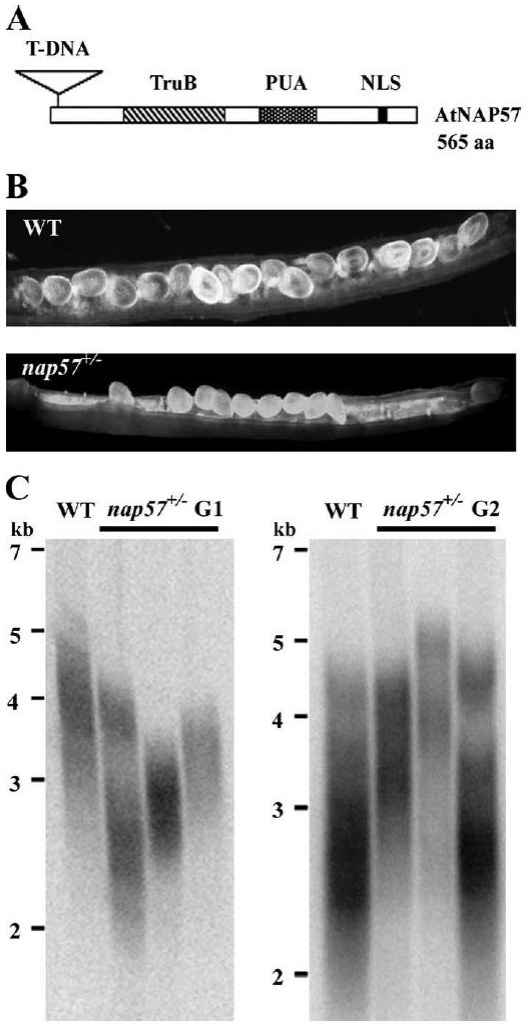


Fig A1-4. FIG. 4. At*NAP57* is an essential gene in *Arabidopsis*. (A) Schematic diagram of the At*NAP57* coding region showing the position of the T-DNA insertion, pseudouridine synthase domain (TruB), pseudouridine synthase, archeosine transglycosylase domain (PUA), and nuclear localization signal (NLS). (B) Siliques (seed pods) from wild-type (WT) or *nap57*/ plants were visualized by microscopy. A reduced seed set was observed for *nap57*/ plants, implying that the homozygous mutation is lethal. (C) TRF analysis of WT, first-generation (G1), or second-generation (G2) *nap57*/ plants. Molecular size markers are indicated.

The T66A mutation in AtNAP57 results in a new shorter telomere length set point.

The T66A missense mutation in dyskerin culminates in DC in humans and is associated with a defect in telomere maintenance (32). Since threonine 66 is conserved in AtNAP57 (Figure A1-5A), we asked whether an alanine substitution at this site would lead to a telomere-related phenotype in plants. Plants heterozygous for the T-DNA insertion in *AtNAP57* were transformed with an *AtNAP57* gene carrying the T66A mutation under the control of its native promoter (Figure A1-5B). We expected to obtain plants heterozygous or homozygous with respect to the T-DNA insertion, and which also carried the exogenous T66A NAP57. Surprisingly, we failed to recover plants that were homozygous for the T-DNA insertion in *AtNAP57* and also contained the T66A transgene. Thus, the *AtNAP57* gene bearing the T66A mutation is unable to rescue plants homozygous for the T-DNA insertion.

To test the effect of T66A NAP57 on bulk telomere length, TRF analysis was performed on first generation of transformants expressing T66A NAP57 (T1). Based on previous transformation experiments (Figure A1-2; refs. (42, 49)), we expected that any detrimental consequences of the T66A mutation would be evident in a population of 10-20 transformants (T1 generation). Accordingly, we examined twenty independent transgenic lines. While for most plants bulk telomeres were in the wild-type range (Figure A1-5C, lanes 2 and 3; data not shown), a subset of telomeres in several plants were significantly shorter than wild-type and their shortest telomere tracts trailed down to below 1.6 kb in length (Figure A1-5C, lanes 4 and 5). To determine whether telomeres would continue to shorten in subsequent generations, we monitored the

progeny of one T1 plant (Figure A1-5C, lane 4; Figure A1-5D). For all of the T2 progeny, the shortest telomere tracts migrated below 2 kb (Figure A1-5C, lanes 7-10), and for one plant, the range of telomere lengths was nearly identical to its T1 parent, spanning ~1.2 to 3.5 kb (Figure A1-5C, compare lanes 4 and 7). When this T2 plant was propagated to T3, the telomere tracts of all four progeny were of similar lengths to their T2 parent (Figure A1-5C, lanes 12-15; Figure A1-5D). The loss of longer telomeres was more pronounced in T3 plants (Figure A1-5D). The average bulk telomere length in wild-type plants (~ 4 kb) was reduced by 2 kb in T3 T66A mutants to ~2.1 kb (Figure A1-5D). We detected no additional shortening in two T4 plants monitored (data not shown). Furthermore, none of the shortest telomeres fell below 1 kb in the four generations the transgenic plants were propagated (Figure A1-5D). Thus, the T66A mutation perturbs telomere length regulation, but does not result in progressive telomere shortening. Instead, this mutation appears to promote the establishment of a new shorter length set point.

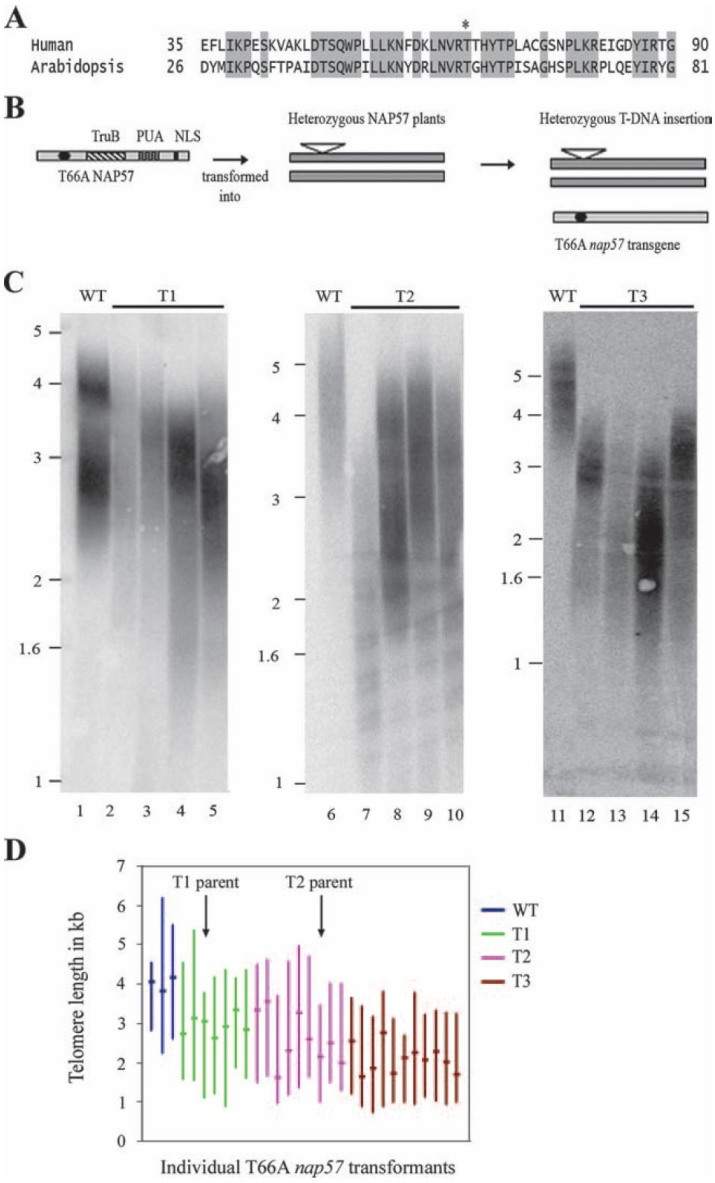


Fig A1-5. The T66A mutation in AtNAP57 results in the establishment of a shorter telomere length set point. (A) Sequence alignment of human dyskerin and AtNAP57 proteins. Conserved residues are highlighted in gray boxes, and the threonine residue targeted for mutagenesis is denoted by an asterisk. (B) Overview of the process for introduction of the T66A mutation in AtNAP57 into *nap57/* plants. (C) TRF analysis of first, second, and third (T1, T2, and T3) generations of T66A transformants. The T1 plant whose telomeres were analyzed in the left panel, lane 4, was used as the parent for T2 progeny plants analyzed in the middle panel. The T2 plant represented in the middle panel, lane 7, was the parent for the T3 progeny analyzed in the right panel. DNA samples were not run as far into the gel shown the right panel as in the other two gels. (D) Graphic representation of bulk telomere length size range and peak telomere length (indicated by -) for WT and T66A transformants is shown. Arrows indicate telomere length measurements for plants used as T1 and T2 parents.

The T66A mutation in AtNAP57 deregulates telomere length on individual chromosome ends.

To further examine telomere length dynamics in T66A NAP57 mutants, we monitored telomeres on three individual chromosome arms: the right arm of chromosome 2 (2R), the left arm of chromosome 3 (3L), and the right arm of chromosome 4 (4R) using Primer Extension Telomere Length Amplification (PETRA). As expected, PETRA produced a single diffuse band for each telomere in wild-type samples (Figure A1-6A, lanes 1-3), consistent with tight regulation of telomere tracts on homologous chromosomes (41). A similar profile was observed in *nap57*$^{+/-}$ mutants (Figure A1-6A, lanes 4-6). A different result was obtained with T66A mutants. In the T1, T2 and T3 mutants that displayed shorter bulk telomere lengths, individual telomere tracts appeared as a cluster of several sharp bands (Figure A1-6A, lanes 10-18; Figure A1-6B). In contrast, the siblings of these plants whose bulk telomeres fell within the wild-type range showed a PETRA profile that more closely resembled to wild-type plants (Figure A1-6A, lanes 7-9; Figure A1-6B). This finding implies that the telomere length regulation on homologous chromosomes is perturbed in T66A *nap57* mutants. As discussed below, this phenotype is consistent with decreased telomerase activity.

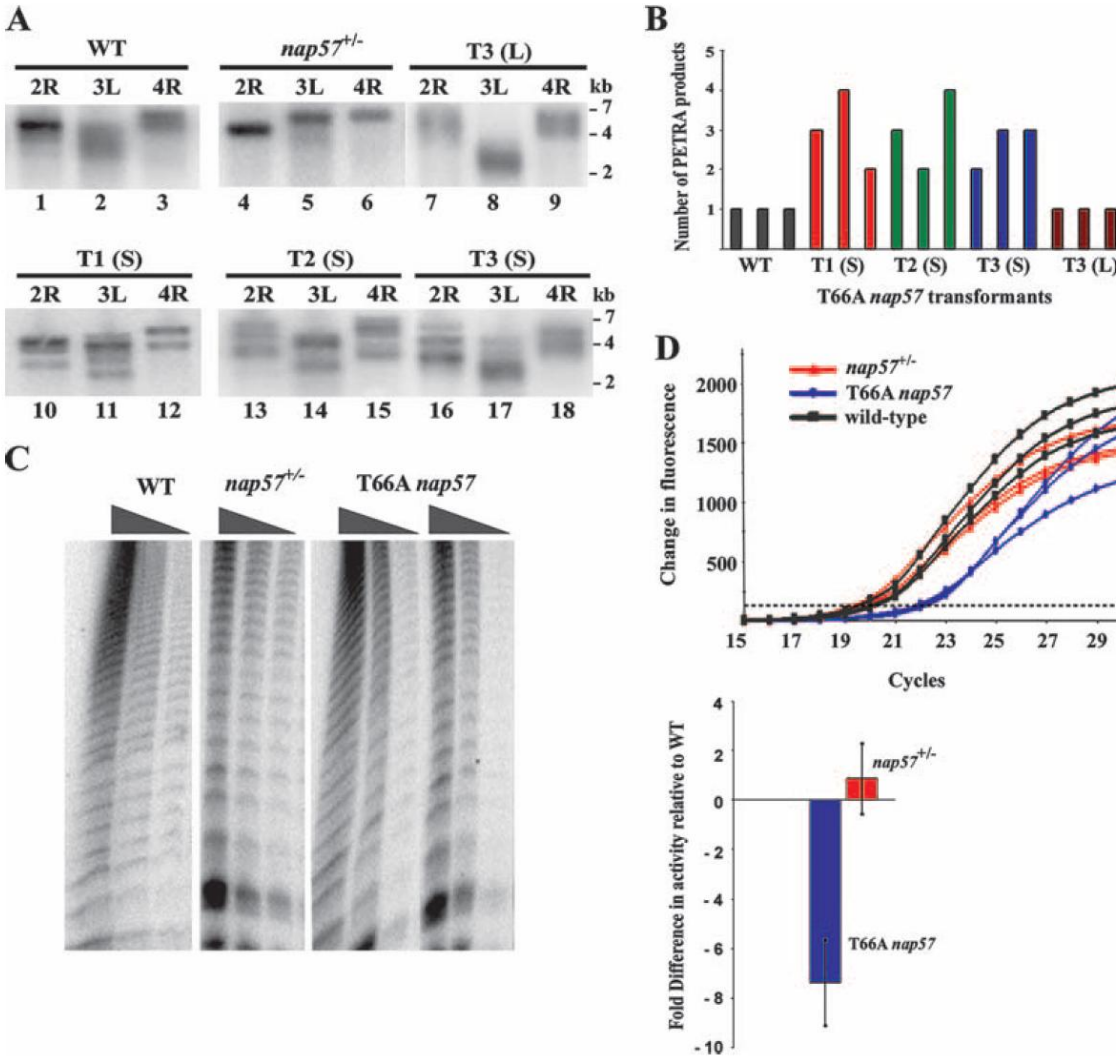


FIG.A1-6. The T66A mutation in AtNAP57 affects telomere length regulation on individual chromosome ends and decreases telomerase activity in vitro. (A) PETRA results are shown for the WT and individual *nap57*/, T1, T2, and T3 T66A *nap57* transformants with short telomeres (S) and a T3 T66A *nap57* transformant with wild-type-length telomeres (L). The telomeres monitored are indicated. 2R, right arm of chromosome 2; 3L, left arm of chromosome 3; 4R, right arm of chromosome 4. (B) Graphic representation of PETRA products obtained in each reaction as determined by visual inspection. (C) TRAP assay results for the WT and *nap57*/ and T66A *nap57* transformants. Reactions were conducted using 1:50, 1:500, and 1:5,000 dilutions of protein extracts. (D) Results of real-time TRAP. The top panel shows raw data for three (each) of the WT and *nap57*/ and T66A *nap57* transformants. The dashed line represents the threshold cycle for TRAP product detection. The bottom panel shows a histogram of the telomerase activity levels for *nap57*/ and T66A transformants relative to those for the WT. Extracts from 10 individual plants from each genotype were monitored.

The T66A mutation in AtNAP57 decreases telomerase activity in vitro and in vivo.

To determine whether the T66A mutation in AtNAP57 decreased telomerase enzyme activity *in vitro*, TRAP assay was performed. Extract titration experiments revealed no detectable difference in the level of telomerase activity in wild-type versus *nap57*$^{+/-}$ mutants (Figure A1-6C), supporting the conclusion that AtNAP57 is not haploinsufficient in *Arabidopsis*. In contrast, TRAP conducted at the highest dilution (1:5000) of protein extract reproducibly revealed decreased *in vitro* telomerase activity in T66A transformants relative to *nap57*$^{+/-}$ siblings (Figure A1-6C). To more precisely gauge the level of telomerase activity in these mutants, we performed quantitative real-time TRAP following a method developed to monitor telomerase activity levels in human cells (22). As expected, we found no significant difference in telomerase activity in extracts prepared from *nap57*$^{+/-}$ versus wild-type plants. In contrast, T66A *nap57* mutants showed a seven-fold decrease in enzyme activity (Figure A1-6D).

To examine the effect of the T66A *nap57* mutation on telomerase activity *in vivo*, we studied the consequences of this mutation in *ku70* mutants. KU70 is best known for its role in the non-homologous end joining (NHEJ) DNA repair pathway (36), but in *Arabidopsis* it also acts as a potent negative regulator of telomerase (13, 38). Telomeres in *ku70* mutants expand to two to three-fold the size of wild-type in a single generation and this elongation is dependent on telomerase (38, 49). Thus, we predicted that incorporation of T66A *nap57* into the telomerase RNP would diminish the enzyme's ability to elongate telomeres in a *ku70* background. To test this hypothesis, we crossed the T66A *nap57* transformants with *ku70*$^{-/-}$ plants (Figure A1-7A). In the first (F1)

generation, we generated plants heterozygous for *AtKU70* and *AtNAP57* and selected for the T66A mutant transgene. In the second generation (F2), we recovered plants that were *ku70*$^{-/-}$ *nap57*$^{+/-}$ and carried the mutant T66A transgene. TRF analysis of the F2 population showed that *ku70*$^{-/-}$ *nap57*$^{+/-}$ plants elongated their telomeres to 5-8 kb (Figure A1-7B, lane 8), while telomeres remained short in the presence of the T66A transgene (Figure A1-7B, lanes 1-7). We conclude that expression of the T66A *nap57* allele prevents telomerase from hyper-elongating telomeres when its negative regulator *KU70* is inactivated. Altogether, our data argue that AtNAP57 is essential for maximal activity of the *Arabidopsis* telomerase RNP *in vivo*.

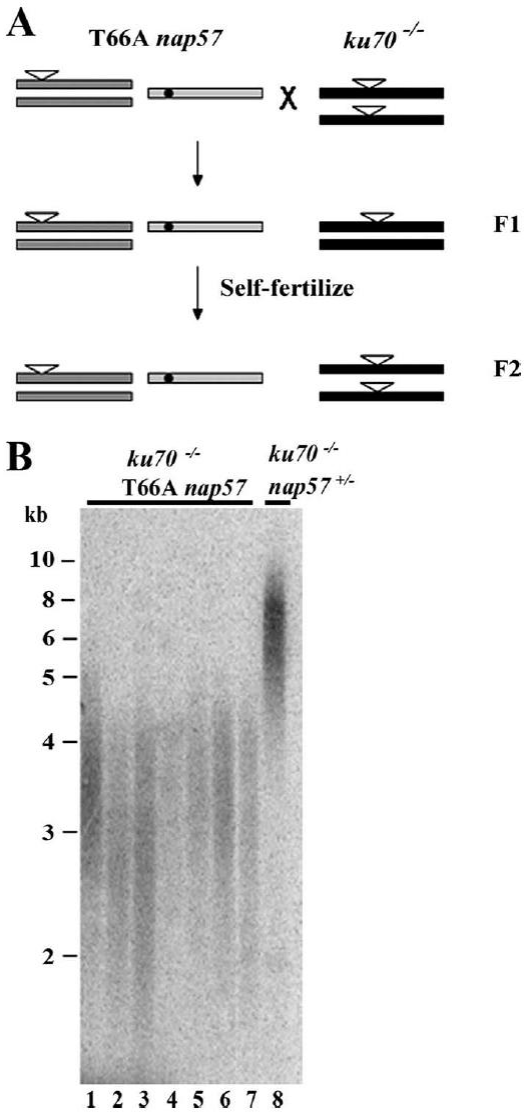


Fig A1-7. The T66A mutation in AtNAP57 reduces telomerase activity in vivo. (A) Schematic diagram of genetic crossing scheme to generate *ku70*/ mutants carrying the T66A *nap57* allele. (B) TRF analysis of seven *ku70*/ T66A *nap57* plants and one *ku70*/ *nap57*/ control plant is shown.

DISCUSSION

A conserved pathway for telomerase biogenesis in higher eukaryotes

The telomerase RNP is evolving at a rapid pace. The TR and TERT subunits have diverged dramatically, and a distinct set of proteins has emerged in higher and lower eukaryotes to promote RNP biogenesis and enzyme action at the chromosome terminus (6, 8). The yeast TR (TLC1) bears a Sm-protein binding motif and has adopted an RNP biogenesis scheme similar to snRNPs (40), while vertebrate TRs have acquired a 3' H/ACA box domain found in snoRNAs and are bound by the dyskerin complex. Thus, although dyskerin's function in catalyzing pseudouridylation of ribosomal RNAs is conserved, in mammals it has evolved an additional, more specialized role as an integral component of the telomerase RNP (7, 32).

In this study we provide several lines of evidence that telomerase enzymes from higher eukaryotes share a requirement for dyskerin. First, we found that AtNAP57 and AtTERT proteins co-localize to the nucleolus in *Arabidopsis*. Nucleolar localization of telomerase could be especially advantageous for *Arabidopsis*, since telomeres cluster at the nucleolar periphery (2). Second, we showed that AtNAP57 physically associates with enzymatically active telomerase particles. The major interaction partner for AtNAP57 in the plant telomerase RNP is likely to be TR, since AtNAP57 association with telomerase is abolished following RNase A treatment. Intriguingly, we also discovered a novel, but weak, interaction for AtNAP57 with AtPOT1a. AtPOT1a is an OB-fold bearing protein that physically interacts with catalytically active telomerase in *Arabidopsis* and promotes enzyme function *in vitro* and *in vivo* (43). Recent studies

indicate that human telomerase associates with TPP1, an hPOT1 binding partner that also harbors a OB-fold motif (48, 50). Whether the human dyskerin contacts TPP1 is unknown. Third, and most importantly, we demonstrated that AtNAP57 is crucial for the function of *Arabidopsis* telomerase. Transgenic plants bearing a mutant AtNAP57 allele display reduced levels of telomerase activity *in vitro* and perturbed telomere length regulation *in vivo* (see below).

Arabidopsis is not haploinsufficient for its known telomerase components.

Essential components of the telomerase RNP are limiting in mammals and yeast. In both *terc*$^{+/-}$ (TR) and *tert*$^{+/-}$ ES mouse cells, telomere maintenance is compromised (10, 17, 27). Indeed, haploinsufficiency of hTR is directly linked to autosomal dominant DC and the reduced hTR levels along with shorter telomeres in these patients results in disease anticipation (46). Similarly, recent studies indicate that TLC1, the TR component in *S. cerevisiae*, is haploinsufficient for telomere maintenance (34). Moreover, yeast heterozygous for both TLC1 and EST1, a telomerase-associated protein, exhibit a phenomenon referred to as additive haploinsufficiency, where telomere tracts are even shorter than in either single heterozygote (24, 26).

In contrast, the known telomerase-associated proteins in *Arabidopsis*, AtTERT (12), AtPOT1a (43), and AtNAP57 (this study) are not haploinsufficient for telomere maintenance. Plants heterozygous for these components display wild-type levels of telomerase activity *in vitro*, and maintain telomeres in the wild-type size range through multiple generations. While it is possible that TR will prove to be present in limiting

quantities, *Arabidopsis* may simply require a very low level of telomerase to maintain genome stability. We note that the *Arabidopsis* genome is comprised of only 10 chromosomes (twenty telomeres), significantly fewer than in diploid budding yeast (64 telomeres) or in human (92 telomeres) cells.

The T66A mutation in AtNAP57 acts as a dominant negative allele to decrease telomerase activity *in vitro* and *in vivo*.

The T66A mutation in human dyskerin leads to DC through reduction in the steady state level of hTR, decreased telomerase activity and progressive telomere shortening (32). To determine if *Arabidopsis* would exhibit similar defects in telomerase, we generated transgenic plants harboring the corresponding mutant allele. As for humans (32), the T66A mutation in AtNAP57 did not grossly affect rRNA processing in *Arabidopsis* (K. Kannan and D.Shippen, unpublished data). Nonetheless, this mutation is highly deleterious; expression of this allele could not rescue the lethality associated with plants lacking both copies of the wild-type AtNAP57. Furthermore, although plants harboring one wild-type copy of *AtNAP57* and the T66A *nap57* transgene were viable, they displayed decreased telomerase activity both *in vitro* and *in vivo*. We suspect that this outcome is a consequence of reduced stability of telomerase RNA, but testing this hypothesis awaits identification of this molecule. Nonetheless, in marked contrast to the fate of human telomeres in T66A DC cells (32), telomeres in T66A *nap57* transgenic plants did not undergo progressive shortening. Instead, telomeres were stably maintained at a length approximately 2 kb shorter than in wild-type.

Data from PETRA on individual chromosome ends argue that this new length set point for telomeres is a consequence of limiting telomerase activity. In the PETRA assay, wild-type telomeres on homologous chromosomes appear as a single heterogeneous band (Figure6A, ref. 41). Because DNA is analyzed from an entire plant, these results mean that individual telomere tracts are subjected to extremely tight regulation during plant growth and development (39). Strikingly, in T66A *nap57* mutants PETRA generates a complex profile of multiple sharp bands, indicating that telomere length is deregulated on individual chromosome ends. Previous studies in yeast (44), mammals (21), and in *Arabidopsis* (41) show that telomerase acts preferentially on the shortest telomeres in the population. We hypothesize that substrate preference is exacerbated in plants with reduced levels of telomerase (e.g. T66A *nap57* mutants), resulting in elongation of only a subset of telomeres in a fraction of cells. Bulk telomeres can then establish a new length set point when equilibrium between the compromised telomerase and forces that shorten telomeres (eg. the end-replication problem, nuclease action and recombination) is attained. Because the demand for telomerase activity is significantly greater in humans, we suspect such cells bearing the T66A *nap57* mutation fail to achieve a new telomere length set point and suffer progressive telomere erosion.

How does the T66A mutation in AtNAP57 inhibit telomerase activity in *Arabidopsis*? Since AtNAP57 is not haploinsufficient for telomerase function in plants, but telomerase is inhibited when the T66A *nap57* allele is introduced into this background, the data argue that T66A NAP57 acts as a dominant negative inhibitor. Catalytically active human telomerase is a 670 kDa dimer comprised of two TERT, two

TR, and two dyskerin molecules (7). *Arabidopsis* telomerase is approximately the same molecular mass and like hTERT (4), AtTERT is capable of dimerization *in vitro* (C. Cifuentes-Rojas, K. Kannan and D. Shippen, unpublished data). Thus, the plant telomerase may also harbor two copies of AtNAP57. Accordingly, incorporation of a mutant form of this protein into the telomerase RNP may compromise enzyme function.

APPENDIX II

PROTECTION OF TELOMERES 1 IS REQUIRED FOR TELOMERE INTEGRITY IN THE MOSS *PHYSCOMITRELLA PATENS**

Summary

In vertebrates, the single-stranded telomeric DNA binding protein Protection of Telomeres 1 (POT1) shields chromosome ends and prevents them from eliciting a DNA damage response. By contrast, *Arabidopsis thaliana* encodes two divergent full-length POT1 paralogs that do not exhibit telomeric DNA binding *in vitro* and have evolved to mediate telomerase regulation instead of chromosome end protection. To further investigate the role of POT1 in plants, we established the moss Physcomitrella patens as a new model for telomere biology and a counterpoint to Arabidopsis. The sequence and architecture of the telomere tract is similar in P. patens and Arabidopsis, but P. patens harbors only a single-copy POT1 gene. Unlike At POT1 proteins, Pp POT1 efficiently bound single-stranded telomeric DNA *in vitro*. Deletion of the P. patens POT1 gene resulted in the rapid onset of severe developmental defects and sterility. Although telomerase activity levels were unperturbed, telomeres were substantially shortened, harbored extended G-overhangs, and engaged in end-to-end fusions. We conclude that the telomere capping function of POT1 is conserved in early diverging land plants but is subsequently lost in Arabidopsis.

---

*Reprinted with permission from Shakirov E.V., Perroud P.F., Nelson A.D.L., Cannel M.E., Quatrano R.S., and Shippen D.E. 2010. Protection of Telomeres 1 is required for telomere integrity in the moss *Physcomitrella patens*. *Plant Cell* **22**, 1838-1848.

INTRODUCTION

Telomeres are nucleoprotein complexes that physically cap the ends of eukaryotic chromosomes. Telomeric DNA promotes genome stability through elaborate interactions with a plethora of telomere-associated proteins. The evolutionarily conserved POT1 (Protection Of Telomeres 1) protein is a key component of shelterin, the vertebrate telomere capping complex (de Lange, 2005). POT1 binds to the single-stranded (ss) G-rich region of the chromosome terminus via structurally conserved N-terminal oligonucleotide/oligosaccharide binding folds (OB-folds) (Lei et al., 2003; Lei et al., 2004). Through interactions with its binding partner TPP1 (Houghtaling et al., 2004; Liu et al., 2004; Ye et al., 2004), POT1 forms a bridge to the double-stranded (ds) region of the telomere tract. Deletion of *POT1* in *Schizosaccharomyces pombe* causes rapid telomere loss and cell death, suggesting that its main function is to protect telomeres from degradation (Baumann and Cech, 2001). In vertebrates, the most critical POT1 function appears to be its ability to block ATR (Ataxia Telangiectasia and Rad3 related)-dependent recognition of telomeric DNA as sites of DNA damage (Denchi and de Lange, 2007; Churikov and Price, 2008). POT1 also plays a role in both telomere length regulation by modulating telomerase activity and in chromosome end protection by preventing nucleolytic attack of the C-rich telomeric DNA strand (reviewed in (de Lange, 2009).

Arabidopsis thaliana is the reference species for plant biology and consequently, telomere research has focused on this model. Although orthologues of several yeast and

vertebrate telomere-related genes have been described (Riha et al., 2001; Song et al., 2008), a technical obstacle for the functional characterization of other candidates is that T-DNA insertion mutants are currently unavailable for a substantial fraction of Arabidopsis genes (O'Malley and Ecker, 2010). Genetic analysis is also complicated by the multiple whole genome duplications associated with Brassicaceae, the family to which Arabidopsis belongs (Beilstein et al, 2010). For example, sequence homologues of the mammalian shelterin subunits TRF1 and TRF2 are encoded by at least six different *TRFL* (TRF-like*)* genes in *A. thaliana*, which appear to be partially redundant (Karamysheva et al., 2004). Furthermore, Arabidopsis encodes three divergent POT1 paralogues, including two full-length proteins, At POT1a and At POT1b, and a truncated isoform termed At POT1c (Shakirov et al., 2005; Rossignol et al., 2007).

This pattern of extensive gene duplication and divergence in Brassicaceae suggests that lessons learned from Arabidopsis may not always be applicable to telomere biology in other plants. Indeed, while POT1 is critical for telomere integrity in yeast and vertebrates, this is not true for At POT1a and At POT1b. Like the other Brassicaceae POT1 proteins examined to date, At POT1a and At POT1b exhibit no detectable binding to ss telomeric DNA *in vitro* (Surovtseva et al., 2007; Shakirov et al., 2009a). Notably, At POT1a physically associates with the telomerase RNP and is critical for telomere length maintenance *in vivo* (Surovtseva et al., 2007). Although initial experiments with over-expression of dominant-negative mutants suggested that At *POT1b* is involved in chromosome end protection (Shakirov et al., 2005), recent analysis of a null allele revealed that At POT1b is a negative regulator of telomerase activity and plays no

significant role in telomere protection (E. Shakirov, A. Nelson and D. Shippen, unpublished results). Thus, At POT1a and At POT1b have evolved to function in the telomerase pathway. To fully elucidate the spectrum of *POT1* functions in the plant kingdom, it will be necessary to examine the contribution of *POT1* in other plant species.

Mosses (bryophytes) belong to the oldest diverging clade of extant land plants, whose ancestors made the first move from aquatic to terrestrial habitats some 450 million years ago (Zimmer et al., 2007). Unlike its younger cousins angiosperms (flowering plants), the moss *Physcomitrella patens* possesses a remarkably efficient system of homologous recombination, which accounts for an unprecedented success rate in targeted gene replacement among higher eukaryotes (Quatrano et al., 2007). Gene targeting in *P. patens* is not only five orders of magnitude more efficient than in angiosperms, but also two orders of magnitude more efficient than in the embryonic stem cells of mice, and comparable with that observed in *Saccharomyces cerevisiae* (Schaefer, 2002). Virtually any non-essential *P. patens* gene can be deleted, and the resulting moss phenotype can be screened within several weeks following transformation.

With its recently sequenced genome (Rensing et al., 2008), and excellent molecular and genetic tools (Quatrano et al., 2007), *P. patens* is poised to become the new "green yeast" of plant biology (Schaefer, 2001). Due to its phylogenetic position within the plant kingdom, *P. patens* also provides a unique opportunity to trace the evolution of telomere-related genes from the first primitive land plants to the most

developmentally advanced angiosperms. Here we describe the initial characterization of telomere biology in *P. patens* and show that this species possesses canonical plant TTTAGGG telomere repeats. However, unlike Arabidopsis, *P. patens* harbors only a single *POT1* gene. Deletion of Pp *POT1* results in catastrophic loss of bulk telomeric DNA, increased G-overhangs and end-to-end chromosome fusions. These findings point to an essential role for Pp *POT1* in chromosome end protection similar to its yeast and vertebrate orthologues, but distinct from Arabidopsis.

METHODS

Moss techniques and generation of *POT1* deletion and complementation constructs

Growth of *Physcomitrella patens* isolates Gransden (Gd) and Villersexel (Vx-1) and protoplast generation and transformation were performed as described previously (Perroud and Quatrano, 2008). The gene deletion construct was designed to remove the full ORF of Pp *POT1* from start codon to stop codon. 1.2 kb of genomic DNA immediately upstream of the start codon was PCR amplified with primers GCATCCTAGGCGTGTGATCCCGCAAT-3' and 5'-GCATCTCGAGGAGAACAGACGATTATGTAAG-3' and inserted 5' to the Hygromycin resistance *HptII* gene in the pBHRF vector (Schaefer et al., 2010), using *Avr*II and *Xho*I restriction enzymes. Similarly, 1.2 kb of genomic DNA immediately downstream of the Pp *POT1* stop codon was PCR amplified with primers 5'-GCATGCGGCCGCGAGCCAATTTTTTTTTTGGCTTTG-3' and 5'-GCATACTAGTATGAAGATAATAGTGC-3' and inserted 3' to the Hygromycin

resistance cassette using *Not*I and *Spe*I restriction enzymes. The final vector containing the Hygromycin resistance cassette, flanked by the targeting *P. patens* genomic DNA sequence on both sides, was cut with *Avr*II and *Spe*I and introduced into wild-type *P. patens* isolate Gd by protoplast transformation (Schaefer and Zryd, 1997). Double crossing-over with this construct results in the complete deletion of *POT1* ORF, and the entire Δ*pot1* locus was subsequently PCR amplified with primers 5'-GCTCATACAACAAGCACATTGAC-3' and 5'-CAACTTCATCCAACCATGCAG-3', flanking the *POT1* locus, and sequenced to verify its absence.

To generate the complementation construct Δ*pot1*-Pp-*POT1*, the entire Pp *POT1* ORF was PCR amplified with 5'-GAATTCTCTTGGAAAGATGCGAGCGGCTGGTCTATGTTGCAGAGGACCATGTCG-3' and 5'-GATCCTCGAGTCATCCCGGAAATCGTGTAC-3' and cloned into the pBNRF vector (Schaefer et al., 2010) upstream of the G418 (kanamycin) resistance gene *NptII* using the same restriction sites. The targeting *P. patens* genomic DNA sequence on both sides of the complementation construct was the same as in pBHRF and contained the native Pp *POT1* promoter. Δ*pot1* strain was transformed as described above. Stable transformants were selected on G418-containing medium and genotyped for the presence of the rescuing construct in the correct location. DNA gel blot analysis was performed with DIG-labeled PCR probes spanning the targeting sequences as described in (Perroud and Quatrano, 2008).

TRF, PETRA, TF-PCR, protein expression in RRL and EMSA

P. patens DNA was extracted as described (Cocciolone and Cone, 1993). TRF analysis was performed using *Tru1*I (Fermentas, Hanover, MD) restriction enzyme and a ^{32}P 5' end-labeled $(T_3AG_3)_4$ oligonucleotide probe (Fitzgerald et al., 1999). For the *Bal*31 exonuclease assay, 100 µg of *P. patens* genomic DNA was incubated with 30 units of *Bal*31 (New England Biolabs) or with H_2O (0 min time point) in 1X *Bal*31 reaction buffer at 30˚C. Equal amounts of sample were removed at specified intervals for 60 min. Reactions were stopped by the addition of 20 mM EGTA and heating to 65˚C for 15 min. DNA in each sample was precipitated with isopropanol and ammonium acetate, followed by *Tru1*I digestion and DNA gel blotting as described above. TF-PCR was performed essentially as described (Heacock et al., 2004) using primers 5'-CACTACATCGCTGGTCAGAAACGA-3' (specific for chromosome arm A, *P. patens* scaffold_2162) and 5'-GAAGGTATGTCATGGCCTCAAAGCT-3' (specific for chromosome arm B, *P. patens* scaffold_229). Primer B can also be substituted with primer B1 (5'-GATTGCACCATCATTGCCATCGCA-3'), which is located on the same scaffold 32 nt closer to the start of the telomeric tract. Two modifications to the TF-PCR protocol were made: primer annealing was performed at 58ºC, and PCR was run for 30 cycles. PETRA was performed for 19 cycles with primers A or B as described (Heacock et al., 2004). Radioactive signals were scanned by a Pharos FX Plus Molecular Imager (Bio-Rad laboratories, Hercules, CA), and the data were analyzed by either IMAGEQUANT software (Molecular Dynamics) or by Quantity One v.4.6.5 software (Bio-Rad). G-overhang analysis and site-directed mutagenesis were performed as described (Song et al., 2008).

EMSA assays

Expression of plant POT1 proteins in rabbit reticulocyte lysate (RRL) and EMSA assays were conducted as described (Shakirov et al., 2009a). Briefly, each reaction (15 µl total volume) contained equal amounts (4 µl) of RRL-translated plant POT1 protein, 0.5 pmol of ^{32}P-labeled telomeric oligonucleotide, 3 µl of 5 x DNA-binding buffer (100 mM Tris–HCl pH 7.8, 250 mM NaCl, 50 mM MgCl, 5 mM EDTA, 5 mM DTT, 25% glycerol) and 1 µl each of Herring sperm DNA (50 µg/ml), DNA oligonucleotide TTAATTAACCCGGGGATCCGGCTTGATCAACGAATGATCC and RNA oligonucleotide $(UUUAGGG)_5$, which were added as nonspecific competitors (2.5 µM each) as described in (Baumann et al., 2002). Reactions were incubated at RT for 15 min and the complexes were separated on 5% polyacrylamide gel (acrylamide: bisacrylamide 29:1) for 2 h at 150 volts in 0.8 x TBE at RT, dried and exposed to PhosphorImager screens. Screens were scanned by a Pharos FX Plus Molecular Imager and signal intensity was quantified by Quantity One v.4.6.5 software.

RACE and RT-PCR conditions

Total RNA was extracted from plant tissue using Tri Reagent solution (Sigma). Reverse transcription was performed using Superscript III reverse transcriptase (Invitrogen), as described (Shakirov et al, 2005). 5' and 3' RACE (FirstChoice® RLM-RACE Kit, Ambion) to deduce the full-length Pp *POT1* cDNA was performed according to the manufacturer's instructions. To analyze Pp *POT1* expression, RT-PCR was

initially performed with control reactions for Ubiquitin gene expression as described (Harries et al., 2005) to normalize for RNA loading, followed by *POT1* RT-PCR with primers 5'-TCGACCCGGGATGTTGCAGAGGACCATGTC-3' and 5'-GATCCTCGAGTCATCCCGGAAATCGTGTAC-3', which span the entire ORF.

Accession Numbers

Sequence data from this article can be found in the Arabidopsis Genome Initiative or GenBank/EMBL databases under the following accession numbers: AY884593 (At POT1a), AY884594 (At POT1b), Q9NUX5 (h POT1), XP_001766873 (Pp Ubiquitin). Pp *POT1* cDNA was deposited into GenBank, accession EU880302.

RESULTS

Telomere length and composition in model moss species.

The length of the telomere tract varies considerably among different species of land plants and green algae, ranging from 0.5 kb in *Chlorella vulgaris* to over 150 kb in tobacco (Fajkus et al., 1995; Higashiyama et al., 1995). In most *A. thaliana* ecotypes, telomeres span 2-5 kb (Shakirov and Shippen, 2004). To evaluate telomere length in *P. patens,* terminal restriction fragment (TRF) analysis was performed using a radioactive probe consisting of four canonical plant TTTAGGG repeats. The most commonly used isolate of *P. patens,* Gransden (Gd), harbored telomere tracts in the range of 0.6 – 3 kb (Figure 1A, lane 2). Telomeres in the second isolate, Villersexel (VX-1), were slightly longer, in the range of 1 – 3.5 kb (Figure 1A, lane 3). Thus, *P. patens* telomeres are approximately two-fold shorter than in *A. thaliana* (Figure 1A, lane 1), but similar in

length to the telomere tracts in another moss species, *Barbula unguiculata*. We verified that TTTAGGG-hybridizing repeats are located at chromosomal ends using *Bal*31 exonuclease. A true telomeric signal should disappear over time, while interstitial cross-hybridizing regions are resistant to enzyme treatment. With the exception of a single interstitial band (indicated by an asterisk), the bulk hybridization signal was lost by 30 minutes of *Bal*31 treatment (Figure 1B), confirming that these *Bal*31-sensitive TTTAGGG-hybridizing repeats are located at chromosome ends.

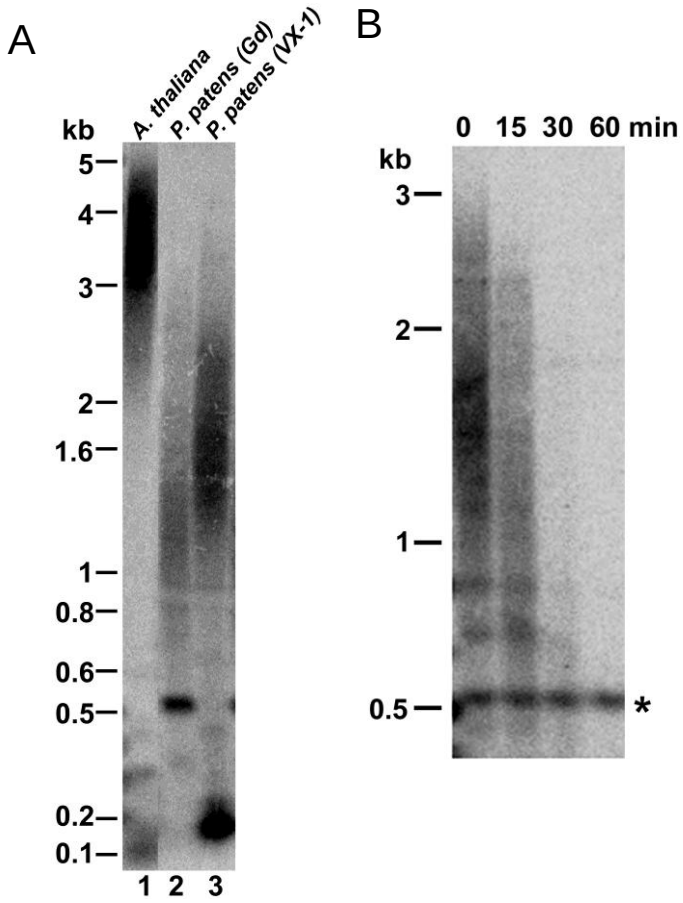


Figure 1. Moss telomere length analysis. (A) TRF analysis of telomeric DNA in *P. patens* isolates Gd (lane 2) and Vx-1 (lane 3). Molecular weight markers are shown on the left. Genomic DNA was digested with *Tru1*I and hybridized with a plant telomere-specific (TTTAGGG)4 probe. *A. thaliana* DNA was used as a control (lane 1). (B) *Bal*31 analysis of genomic DNA from *P. patens* isolate Gransden. *Tru1*I digestion of genomic DNA was performed without prior *Bal*31 treatment (0 min), or after various incubation periods with *Bal*31 exonuclease. Asterisk indicates cross-hybridizing interstitial telomeric DNA bands which is not sensitive to *Bal*31 digestion for up to 60 min.

Identification of *P. patens* POT1

The publicly available *P. patens* genome database (http://moss.nibb.ac.jp/) was searched with the *A. thaliana* POT1a and POT1b sequences as the queries using blastp and tblastn programs. Similar to most other plants analyzed to date (Shakirov et al., 2009b), *P. patens* harbors a single *POT1* gene, which encodes a 497 amino acid protein with 50% and 45% sequence similarity to *A. thaliana* POT1a and POT1b, respectively. The Pp POT1 protein is predicted to contain two N-terminal DNA binding OB-folds with secondary structures similar to the human POT1 protein (Figure 2A). Like At *POT1a* and At *POT1b*, Pp *POT1* consists of ten exons with conserved exon-intron junctions, indicating that *P. patens* and *A. thaliana POT1* genes are indeed orthologous. As noted for other *P. patens* genes (Rensing et al., 2005), Pp *POT1* introns are on average twice as long as in *A. thaliana*. RT-PCR revealed that Pp *POT1* is continuously expressed during gametophytic development (protonemata and gametophore tissues) (Figure 2B, lanes 1 and 2).

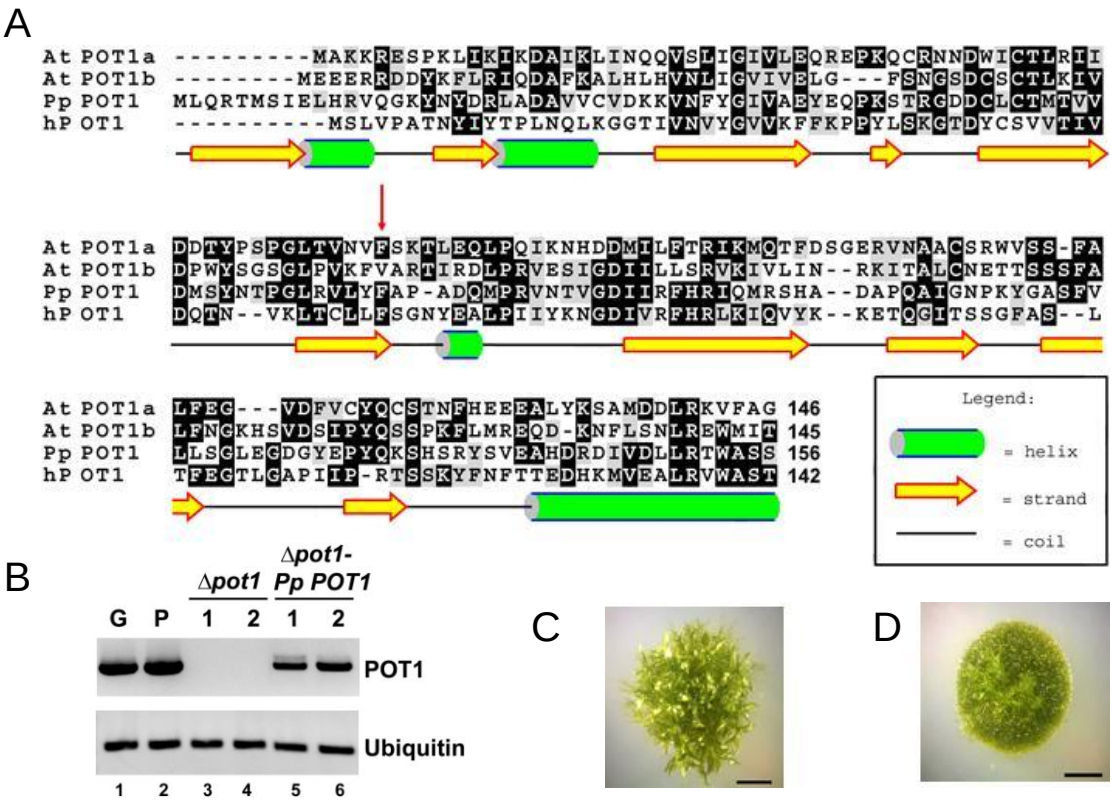


Figure A2-2. Identification and analysis of *Physcomitrella patens* POT1. (A) Amino acid alignment of the predicted PpPOT1 OB1 region with *A. thaliana* (AtPOT1a and AtPOT1b) and human POT1 proteins. The secondary structure of PpPOT1 OB1 was predicted with PsiPred software (McGuffin et al., 2000). Numbers indicate amino acid positions relative to the start codon. Red arrow indicates the position of a biochemically important F62 residue in human POT1 and the corresponding amino acid F74 in PpPOT1. Alignment was generated with MEGA 3 software (Kumar et al., 2004) and visualized in the BOXSHADE format. (B) RT–PCR results of the *PpPOT1* gene expression (top panel) in wild-type (lanes 1, 2), Δ*pot1* (lanes 3, 4) and Δ*pot1-PpPOT1* complemented line (lanes 5, 6). Ubiquitin (bottom panel) was used to normalize for RNA loading. (C) and (D) General morphology of wild-type (C) and Δ*pot1* (D) colonies at the four-week old stage. The wild-type colony harbors multiple leafy-like gametophores, but only filamentous protonemata tissue is visible in Δ*pot1.*

Characterization of Pp POT1 DNA binding activity

Although POT1 proteins from Brassicaceae have thus far failed to bind ss telomeric DNA *in vitro* (Surovtseva et al., 2007; Shakirov et al., 2009a), a recent survey revealed that recombinant POT1 proteins from asparagus, maize and the green alga *Ostreococcus lucimarinus* are capable of efficient telomeric DNA binding (Shakirov et al., 2009b). To examine the nucleic acid binding properties of Pp POT1, we expressed Pp POT1 in rabbit reticulocyte lysate (RRL) and performed electrophoretic mobility shift assays (EMSA) with radioactively labeled cocktails of telomeric oligonucleotides containing seven different permutations of either two (2TELO) or three (3TELO) plant TTTAGGG repeats (Figure 3A). *Asparagus officinalis* POT1 (Ao POT1) served as a positive control (Figure 3A, lane 3) (Shakirov et al., 2009b). While no binding of Pp POT1 to the 2TELO probe was observed (Figure 3A, lane 2), Pp POT1 efficiently bound the 3TELO cocktail (Figure 3A, lane 5). As expected (Surovtseva et al., 2007; Shakirov et al., 2009a), POT1 proteins from Arabidopsis and several other vascular plants showed no detectable telomeric DNA binding (Supplemental Figure 1). Of the seven individual permutations in the 3TELO cocktail probe, Pp POT1 bound $(AGGGTTT)_3$ most efficiently (Supplemental Figure 2).

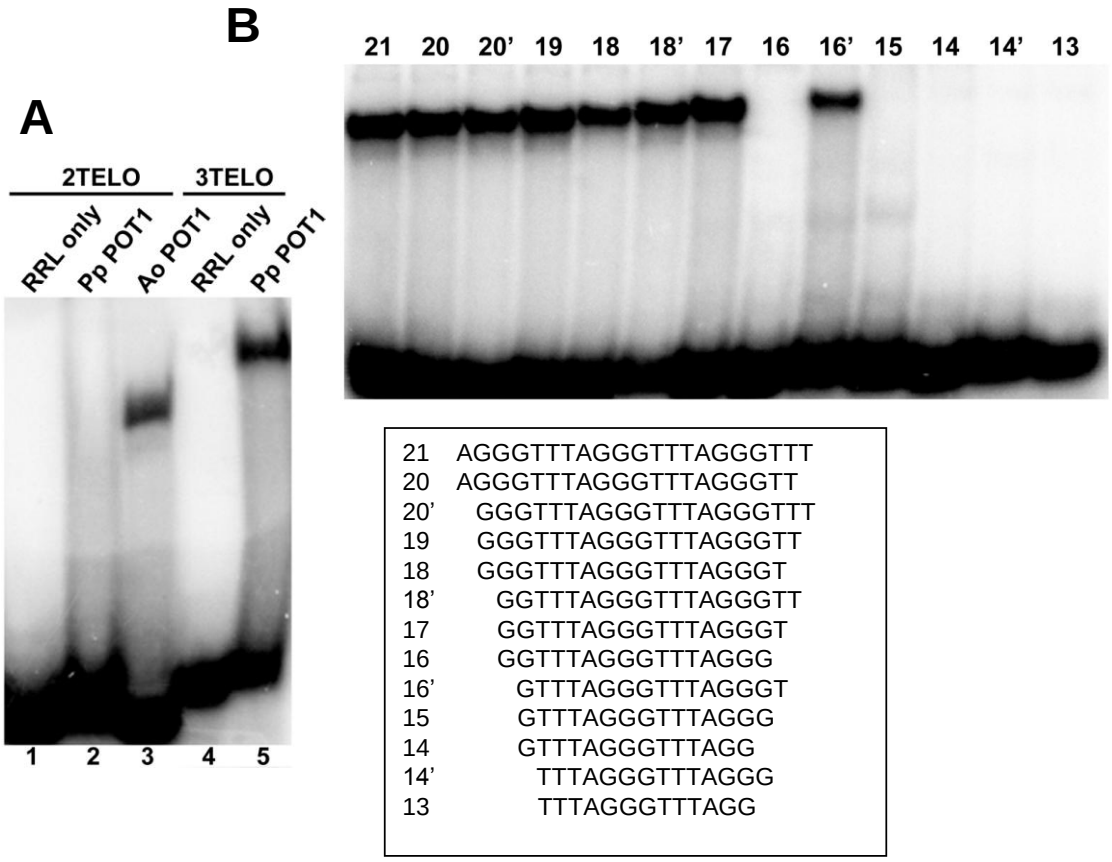


Figure A2-3. Characterization of PpPOT1 interaction with telomeric DNA *in vitro*. (A) EMSA was performed with a cocktail of seven 32P-labeled oligonucleotides corresponding to two (2TELO) (lanes 1-3) or three (3TELO) (lanes 4-5) TTTAGGG repeats. Unprogrammed RRL reactions lacking external DNA template were used as negative controls (lanes 1, 4), and a reaction with asparagus POT1 protein (AoPOT1) was performed as a positive control (lane 3). PpPOT1 binds 3TELO probe (lane 5), but not 2TELO probe (lane 2). (B) Identification of the PpPOT1 minimum DNA binding site. Equal amounts of RRL-expressed PpPOT1 were incubated with the indicated radioactively labeled oligonucleotides (bottom panel) and protein-DNA complexes were separated by native PAGE.

The minimum DNA binding site (MBS) in vertebrate POT1 proteins is highly conserved and consists of 10-12 nucleotides (Lei et al., 2004; Loayza et al., 2004; Wei and Price, 2004). To define the Pp POT1 MBS, we performed EMSA with $(AGGGTTT)_3$ and a series of single nucleotide truncations from either the 5' or 3' end of this substrate (Figure 3B). Removal of the first two nucleotides from both ends of the substrate (oligonucleotide 17) did not decrease binding. However, deletion of one additional nucleotide from the 3' end, but not from the 5' end of this substrate, completely abolished binding (Figure 3B, compare oligonucleotides 16 and 16'). Since Pp POT1 does not bind to two different oligonucleotides containing 15 nucleotides each (Supplemental Figure 3A), we conclude that its minimum tight-binding substrate is the sixteen nucleotide sequence, GTTTAGGGTTTAGGGT.

We next asked which domain(s) in the *P. patens* POT1 are required for MBS recognition. Like POT1 proteins from vertebrates and several plants (Lei et al., 2004; Wu et al., 2006; Shakirov et al., 2009b), both N-terminal OB-folds were necessary for telomeric DNA binding (Supplemental Figure 3B, lanes 4-7). Interestingly, the EMSA profile changed from a single band to a smear in the absence of Pp POT1 C-terminus (Supplemental Figure 3B, compare lanes 2 and 3), suggesting that this region (amino acids 323 – 497) modulates Pp POT1 interaction with DNA *in vitro*.

In mammalian and fission yeast POT1 proteins, an invariant phenylalanine in the first OB-fold (F62 in human POT1, Figure 2A) is required for efficient DNA binding (Lei et al., 2004; Wu et al., 2006). To test if the corresponding residue (F74) is required for telomeric DNA interaction by Pp POT1, we performed EMSA with an F74A Pp

POT1 mutant. The mutant protein failed to bind telomeric DNA *in vitro* (Supplemental Figure 3B, lane 1). Altogether, these data indicate that *P. patens* POT1 is a *bona fide* single-strand telomeric DNA binding protein with structural and biochemical properties similar to POT1 proteins from yeast and vertebrates, but distinct from Arabidopsis.

Morphological and developmental defects in the Δpot1 strain

We generated a deletion of *P. patens POT1 (Δpot1)* by replacing its complete open reading frame (ORF) with the hygromycin-resistance gene *HptII* (Supplemental Figure 4A). The absence of the *POT1* ORF in hygromycin-resistant transformants was verified by PCR genotyping (Supplemental Figure 4B), DNA gel blot analysis (Supplem. Figure 4C), and by sequencing across the entire Δ*pot1* locus. RT-PCR confirmed the absence of *POT1* mRNA in Δ*pot1* (Figure 2B, lanes 3-4). Hereafter we refer to Δ*pot1* as a null mutant.

P. patens was able to grow vegetatively in the absence of *POT1*, suggesting that this gene is not essential for long-term moss survival. Nevertheless, the Δ*pot1* strain displayed a striking loss of developmental program. Although the juvenile protonemata filaments initially grew efficiently, aberrant morphological defects emerged and worsened progressively as the Δ*pot1* advanced to more mature stages of development, ultimately culminating in complete sterility. The growth of gametophores (leafy-like tissue on top of the colony) was significantly delayed in the Δ*pot1* strain (compare Figures 2C and 2D), but the overall number of emerging gametophores was reduced only slightly (Supplemental Figure 5). Furthermore, upon cold treatment, which in wild-type *P. patens* induces the formation of sex organs, the female archegonium and the male antheridium never formed in Δ*pot1*, explaining the sterility phenotype. These data indicate that *POT1* is necessary for the normal development of *P. patens*.

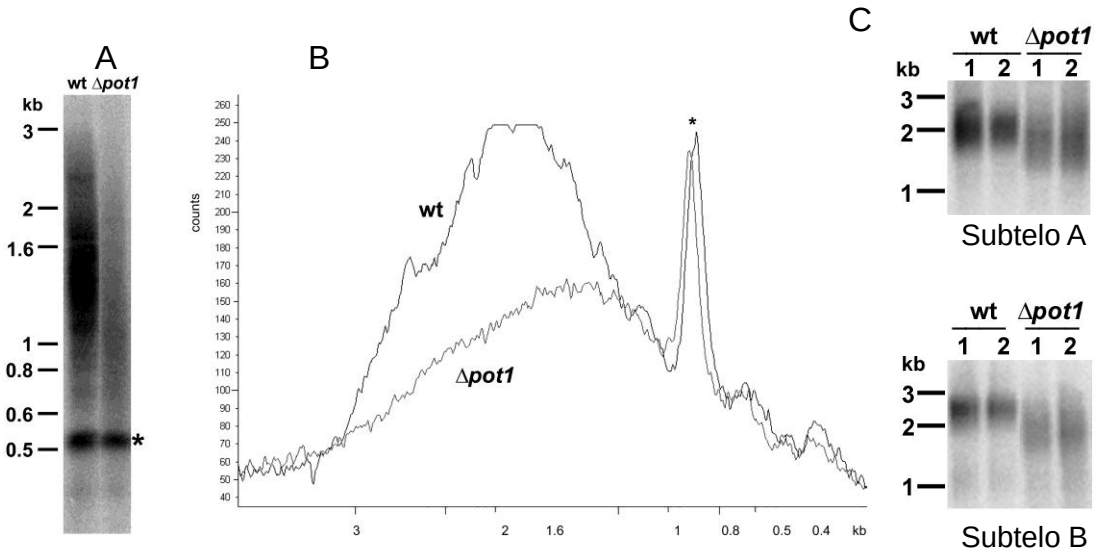


Figure A2-4. Defects in telomere length regulation in Δ*pot1*. (A) TRF analysis of DNA from wild-type (wt) and Δ*pot1*. The asterisk indicates an interstitial cross-hybridizing band used as a loading control. (B) Quantification of results in (A) with ImageQuant software. (C) PETRA results of telomere lengths on two individual chromosome arms, designated A and B, in wild-type (wt) and Δ*pot1* moss. A telomeric probe (TTTAGGG)4 was used to detect TRF and PETRA signals. Molecular weight markers are shown on the left of each panel.

Telomere length maintenance defects in Δpot1

We next asked if the developmental and morphological defects in Δ*pot1* are accompanied by deficiencies in telomere maintenance. TRF analysis revealed a dramatic decrease (~72%) in the overall intensity of telomeric hybridization signal in Δ*pot1* relative to wild-type, indicating that the majority of chromosome ends experienced a massive loss of telomeric DNA (Figure 4A). The average length of the remaining telomere tracts was reduced by ~200 bp (Figure 4B). *Bal*31 digestion of Δ*pot1* DNA confirmed that the residual TTTAGGG-hybridization signal was terminally located (Supplemental Figure 6). Complementation with a wild-type Pp *POT1* gene (hereafter termed Δ*pot1*-Pp-*POT1)* restored telomere length to wild-type, demonstrating that telomere loss was the result of *POT1* deletion (Supplemental Figure 7A and 8A). As expected, RT-PCR confirmed that *POT1* mRNA expression was restored in the Δ*pot1*-Pp-*POT1* line (Figure 2B, lanes 5, 6). We conclude that POT1 is necessary for proper telomere length maintenance in *P. patens*.

Notably, the Δ*pot1*-Pp-*POT1* construct did not complement Δ*pot1* morphological and developmental defects. Although PCR genotyping of Δ*pot1*-Pp-*POT1* strain confirmed the presence of the *POT1* gene in its original locus (Supplemental Figure 4B), DNA gel blot analysis indicated that additional copies of the complementation construct are present in the Δ*pot1*-Pp-*POT1* strain (Supplemental Figure 7B). Thus, increased gene dosage may explain the persistence of developmental defects. Alternatively, the severe genome instability that arises from telomere uncapping (see below) may have resulted in irreversible aberrations in cellular pathways controlling moss morphology and

development, as previously reported for Arabidopsis mutants with dysfunctional telomeres (Riha et al., 2001; Song et al., 2008; Surovtseva et al., 2009).

The TRF assay measures bulk telomere length changes. To explore the dynamics of individual telomere tracts in Δ*pot1*, we employed the primer extension telomere repeat amplification (PETRA) assay originally developed to study Arabidopsis telomeres (Heacock et al. 2004). PETRA amplifies individual telomeric DNA tracts using PCR primers directed at the G-overhang and a unique subtelomeric sequence. Although the sequenced *P. patens* genome is not yet assembled into its 27 chromosomes, two unique subtelomeric regions (arbitrarily designated A and B, see Materials and Methods) were identified. PETRA reactions conducted with wild-type *P. patens* generated a smear spanning up to 0.5kb (Figure 4C), but in Δ*pot1* both A and B telomeres were over 200 bp shorter (Figure 4C). Importantly, the wild-type PETRA profile was restored in Δ*pot1*-Pp-*POT1* (Supplemental Figure 8B).

Telomerase activity is reduced by approximately 10-fold in Arabidopsis plants null for At *POT1a* (Surovtseva et al., 2007). To determine if *in vitro* telomerase activity is altered in Δ*pot1*, we performed a quantitative telomere repeat amplification protocol Q-TRAP (Kannan et al., 2008) with protein extracts from wild-type and Δ*pot1* strains. No statistically significant difference in enzyme activity was observed (Supplemental Figure 9). We conclude that POT1 proteins make fundamentally different contributions to telomere maintenance in *A. thaliana* and *P. patens*.

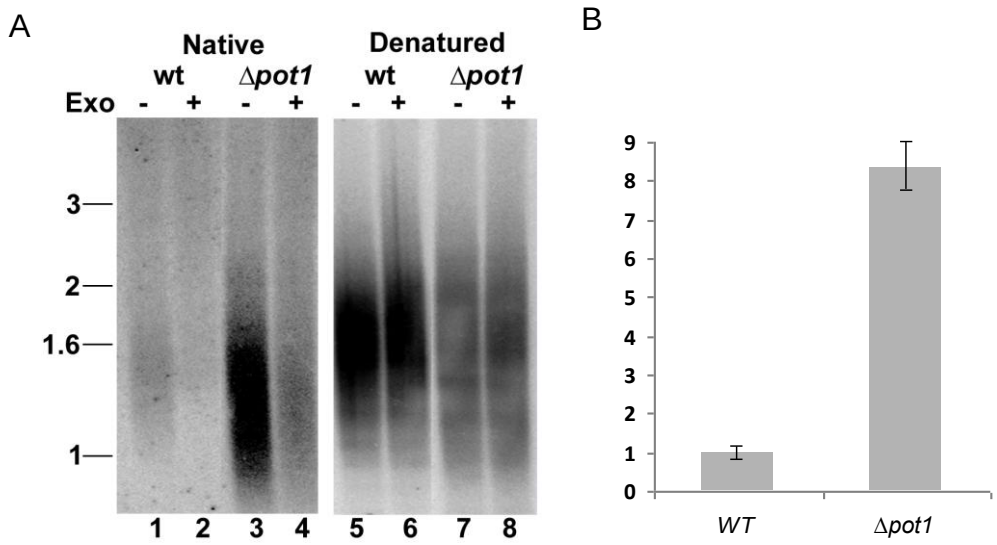


Figure A2-5. Increased G-overhang signals in Δ*pot1* mutants. (A) In-gel hybridization results using a G-strand specific probe (C3TA3)4 to detect telomeric DNA under native (lanes 1-4) and denaturing (lanes 5-8) conditions in wild-type (wt) and Δ*pot1*. T4 DNA Polymerase (a 3'→5' exonuclease) was added to reactions shown in lanes designated by (+). (B) Quantification of the G-overhang signal in (A). Δ*pot1* signal intensity relative to wt is shown (error bars measured for n=4, SD = +/- 0.835).

Defects in telomere architecture and genome stability in Δ*pot1*

Mutations in telomere capping proteins can lead to increased G-overhang length, resulting from increased nucleolytic attack on the C-rich telomeric strand or uncoupled replication of the G-rich and C-rich strands (Linger and Price, 2009). We analyzed G-overhang status in Δ*pot1* mutants using the in-gel hybridization assay. DNA gel blotting with a telomeric DNA probe was performed under native conditions to measure ss telomeric DNA. The gel was re-hybridized with the same probe under denaturing conditions to normalize for the total amount of ds and ss telomeric DNA loaded. Relative to wild-type, the G-overhang signal in Δ*pot1* increased by ~8-fold (Figure 5A, lanes 1 and 3, and Figure 5B). Exonuclease treatment confirmed that hybridization was associated with the chromosome terminus (Figure 5A, lanes 2 and 4). We conclude that *P. patens POT1* is critical for proper chromosome end structure.

S. pombe and mammalian POT1 proteins are required to protect chromosomes from end-to-end fusions (Baumann and Cech, 2001; Veldman et al., 2004; Hockemeyer et al., 2006). By contrast, Arabidopsis *POT1a* and *POT1b* null mutants do not exhibit telomere fusions (Surovtseva et al., 2007); E. Shakirov, A. Nelson and D. Shippen, unpublished results). To evaluate chromosome stability in Δ*pot1*, we employed the Telomere Fusion PCR assay (TF-PCR) (Heacock et al., 2004). As with PETRA, TF-PCR utilizes unique subtelomere-specific PCR primers directed toward the chromosome terminus. If the target chromosome ends are covalently fused, the fusion junction will be amplified and detected by Southern analysis. As expected, no TF-PCR products were observed with DNA samples from wild-type *P. patens* or in the complemented Δ*pot1*-

Pp-*POT1* strain (Supplemental Figure 8C). However TF-PCR products were abundant in reactions with Δ*pot1* DNA (Figure 6A). Interestingly, chromosome fusions were also detected in Δ*pot1* PCR reactions containing only one subtelomere-specific primer (Figure 6B). Since *P. patens* spends most of its lifecycle in the haploid form, these data suggest that sister chromatids fuse in the absence of POT1.

Chromosome end maintenance defects worsen over time in mouse models deficient for either *mPOT1a* or *mPOT1b* (Wu et al., 2006; Hockemeyer et al., 2008). In contrast, we found no evidence of progressive telomere shortening or increased incidence of telomere fusions in long-term vegetatively growing Δ*pot1* cultures. However, since relatively few cultured *P. patens* cells are proliferating (Cove et al., 2006), it is possible that the telomere deprotection phenotype would be exacerbated if the majority of Δ*pot1* cells were forced to divide. To investigate this possibility, we subjected Δ*pot1* cells to several consecutive rounds of colony re-growth from a single progenitor cell (protoplast) (Figure 6C). As predicted, the abundance of TF-PCR products was elevated in Δ*pot1* cultures as the number of protoplasting and new colony formation events increased (Figure 6C). Taken together, our data indicate that POT1 is essential for telomere integrity and chromosome end protection in *P. patens*.

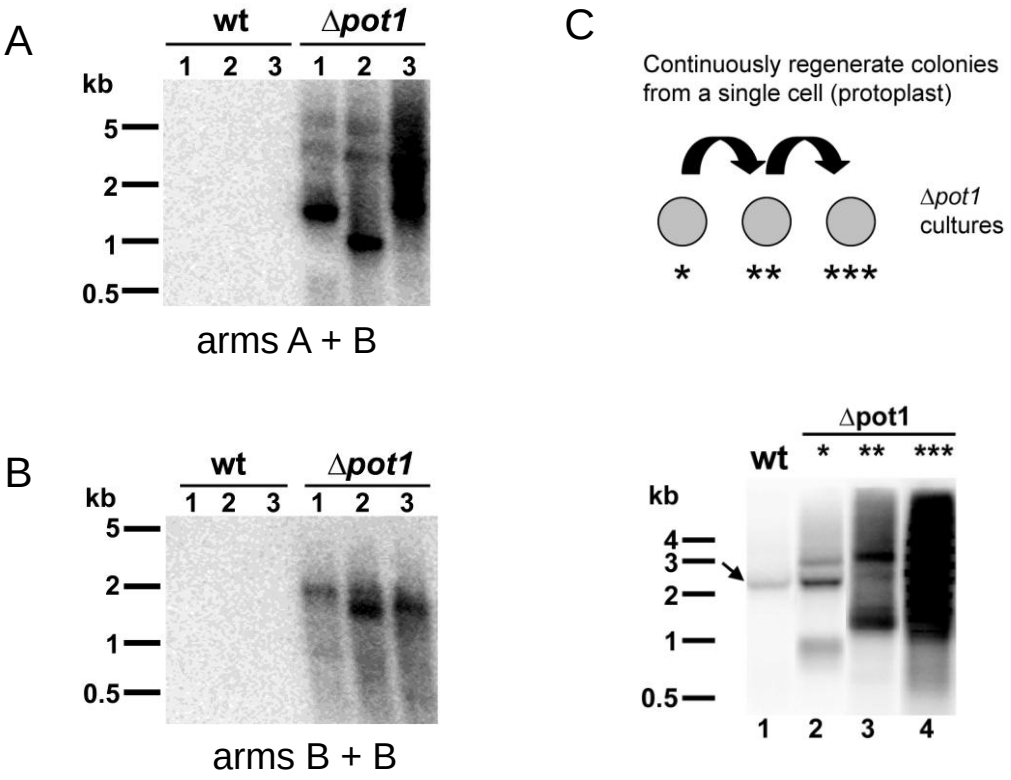


Figure A2-6. Accumulation of chromosome fusions in Δ*pot1*. (A) and (B) TF-PCR detects multiple instances of end-to-end chromosome fusions in Δ*pot1* moss, but not in wild-type (wt). (A) TF-PCR results with subtelomeric primers specific for chromosome arms A and B and DNA from three individual colonies of wt or Δ*pot1*. (B) TF-PCR results with one subtelomeric primer specific for chromosome arm B. (C) Increased incidence of chromosome fusions in Δ*pot1* cultures continuously re-grown from a single protoplast. (Top panel) Diagram of the experimental design. (Bottom panel) TF-PCR results with DNA from wt (lane 1) and Δ*pot1* colonies (lanes 2-4). Δ*pot1* colonies were re-generated from a single cell through protoplasting once (lane 2), twice (lane 3) or three times (lane 4) (see asterisks in top panel). Arrow indicates a non-specific PCR product detected in some wt samples. Molecular weight markers are shown on the left of each panel.

DISCUSSION

Physcomitrella patens, a new model system for plant telomere research

The ability of eukaryotic cells to distinguish between ds DNA breaks and natural chromosome termini is essential to avoid aberrant genome rearrangements and to ensure continued cell proliferation. The factors responsible for telomere stability are evolving rapidly (Linger and Price, 2009). In vertebrates shelterin represses ATR- and ATM-dependent DNA damage pathways, preventing telomere-specific homologous recombination and non-homologous end-joining reactions (de Lange, 2009). While only a subset of the shelterin components can be discerned in plant genomes, ss telomeric DNA binding proteins are an essential component of the protective telomere cap in plants (Surovtseva et al., 2009) and many other model organisms (Gao et al., 2007; Raices et al., 2008; Linger and Price, 2009; Miyake et al., 2009). POT1 orthologs have been described in a wide variety of eukaryotic lineages, from ciliated protozoa and yeast to humans (Baumann and Cech, 2001; Baumann et al., 2002; Jacob et al., 2007). Despite substantial amino acid sequence conservation, POT1a and POT1b from *A. thaliana* appear to have lost this critical chromosome capping function, raising the possibility that POT1 plays a fundamentally different role in the plant kingdom. The analysis of POT1 in the moss *Physcomitrella patens* reported here rejects this hypothesis and indicates that in early land plant lineages, such as bryophytes, POT1 functions in a similar fashion to its fungal and vertebrate counterparts.

P. patens has the potential to emerge as a powerful new model for telomere biology. Like *A. thaliana*, *P. patens* has short telomere tracts abutted by unique

sequences on at least some chromosome ends. These characteristics allow detection of subtle defects in telomere maintenance and chromosome end protection using the highly sensitive PCR assays of PETRA and TF-PCR (Heacock et al., 2004). Like *A. thaliana*, *P. patens* is genetically tractable, but it has added advantage of a streamlined genome with much less redundancy. Besides a single-copy *POT1* gene, *P. patens* harbors only two *TRFL* genes (as opposed to six in *A. thaliana*). Finally, *P. patens* has an exceptionally high rate of homologous recombination and we exploited this property for targeted deletion of *POT1*. Our results not only provide direct genetic evidence of an evolutionarily conserved role for POT1 in chromosome end protection in plants, but also they demonstrate that *P. patens* can serve as a critical counterpoint to Arabidopsis, facilitating a deeper understanding of plant telomere composition and evolution.

Conservation of POT1 functions in eukaryotes

Although the functions of POT1 differ between *A. thaliana* and *P. patens*, together they encompass several of the roles previously ascribed to POT1 orthologues in mammals and fission yeast. Several studies indicated that mammalian POT1 proteins modulate telomerase action *in vitro* (Wang et al., 2007; Latrick and Cech, 2010; Zaug et al., 2010) and *in vivo* (Colgin et al., 2003; Loayza and de Lange, 2003). Although the details of how POT1 influences telomerase are likely to be organism-specific, this function is conserved in plants, as illustrated by the phenotypes of *A. thaliana POT1a*-deficient plants (Surovtseva et al., 2007). Similar to the budding yeast *ever-shorter-telomere* (*est*) mutants (Lundblad and Szostak, 1989), Arabidopsis *POT1a* null mutants

display progressive telomere shortening at a rate of approximately 200-500 bp per plant generation (Surovtseva et al., 2007). As previously described for plants lacking telomerase catalytic subunit TERT (Riha et al., 2001), this gradual telomere erosion leads to a delayed onset of developmental and cell proliferation defects apparent only after multiple plant generations. In contrast, *P. patens* Δ*pot1* mutants exhibit a rapid onset of telomere uncapping characterized by extensive telomere shortening, increased G-overhang length and chromosome fusions. These phenotypes are reminiscent of the situation in *Schizosaccharomyces pombe Pot1*-deficient cells, where chromosome ends rapidly lose all of their telomeric DNA and undergo chromosome circularization (Baumann and Cech, 2001). Thus, the function of *P. patens* POT1 is consistent with an essential role in chromosome end protection.

Like the surviving POT1-deficient *S. pombe* cells, which can not undergo meiosis (Baumann and Cech, 2001), *P. patens* Δ*pot1* cultures fail to undergo sexual reproduction and display a partial loss of developmental program. However, unlike *S. pombe, P. patens* Δ*pot1* mutants do not lose all telomeric DNA and instead telomeres stabilize at a new, shorter length. This outcome may reflect the establishment of a new equilibrium between telomere-shortening activities and extension by telomerase. It is also possible that homologous recombination contributes to the maintenance of telomere tracts in Δ*pot1 P. patens*, as proposed for telomerase-deficient Arabidopsis mutants (Watson and Shippen, 2007). Mouse POT1 proteins prevent recombination of telomeric DNA (He et al., 2006; Wu et al., 2006; Palm et al., 2009), and considering the naturally

high rate of homologous recombination in *P. patens*, this alternative mode of telomere maintenance is certainly feasible.

Although we could detect chromosome fusions in long-term dividing *P. patens* Δ*pot1* cultures, accumulation of these aberrant structures was slow, required multiple cell divisions, and did not result in complete senescence. A similar phenotype is associated with *POT1* deletion in chicken cells (Churikov et al., 2006), implying that some aspects of Pp POT1 function are more similar to vertebrate POT1 than to the *S. pombe* POT1. It is also possible that *P. patens* POT1 functions redundantly with another telomere protection factor, such as a homologue of the putative Arabidopsis ds telomere binding protein TBP1 (Hong et al., 2007) or the newly identified plant CST complex (Surovtseva et al., 2009).

Evolution of POT1 function in plants

Our data underscore remarkable divergence of POT1 functions in lower versus higher plants and provide a framework for elucidating evolutionary changes responsible for the switch from telomere protection to telomerase regulation. We speculate that much of the functional divergence is reflected by changes in the nucleic acid binding properties of plant POT1 proteins. Only a subset of the plant POT1 proteins examined to date display ss telomeric DNA binding activity *in vitro* (Surovtseva et al., 2007; Shakirov et al., 2009a; Shakirov et al., 2009b). Among those that do bind telomeric DNA, significant variation exists within their respective minimal DNA binding site. For example, while Pp POT1 and Ol POT1 (from green alga *O. lucimarinus*) prefer

telomeric DNA substrates terminating with a T on the 3' end, POT1 proteins from the flowering plants prefer oligonucleotides terminating in a G (Shakirov et al., 2009b). The length of the POT1 MBS also varies dramatically among different plant species, ranging from seven to 16 nucleotides (Shakirov et al., 2009b); this study). At 16 nucleotides, *P. patens* POT1 has a very long MBS (Croy and Wuttke, 2006; Shakirov et al., 2009b). The C-terminus of Pp POT1 enhances telomeric DNA binding. It is possible that Pp POT1 oligomerizes through its C-terminal domain, accounting for its unusually long MBS. Notably, the *Asparagus officinalis* POT1, whose MBS is only nine nucleotides, does not require the C-terminal domain for efficient DNA binding *in vitro* (Shakirov et al., 2009b). Alternatively, the extended minimum DNA binding site in Pp POT1 might reflect the contribution of an additional OB-fold in the C-terminus, as has been postulated in yeast Cdc13 and vertebrate POT1 (Theobald and Wuttke, 2004; Wei and Price, 2004).

The most compelling evidence of the rapid evolution of POT1 is illustrated by the Arabidopsis POT1 proteins. Recent data indicate that At POT1a and At POT1b physically interact with telomerase ribonucleoprotein complexes instead of telomeric DNA (Surovtseva et al., 2007)(Cifuentes-Rojas et al., submitted). Thus, we speculate that the switch in plant POT1 proteins from binding telomeric DNA to telomerase fueled their evolution from telomere capping proteins into telomerase regulatory factors. Further analysis of POT1 in Arabidopsis and Physcomitrella may more clearly elucidate the dynamic and evolving interactions between the chromosome terminus and the telomerase enzyme.

APPENDIX III

ATR COOPERATES WITH CTC1 AND STN1 TO MAINTAIN TELOMERES AND GENOME INTEGRITY IN ARABIDOPSIS*

Summary

The CST (CTC1/STN1/TEN1) complex is an essential constituent of plant and vertebrate telomeres. Here we show that CST and ATR act synergistically to maintain telomere length and genome stability in Arabidopsis. Inactivation of ATR, but not ATM, temporarily rescued severe morphological phenotypes associated with ctc1 or stn1. Unexpectedly, telomere shortening accelerated in plants lacking CST and ATR. In first generation (G1) ctc1 atr mutants, enhanced telomere attrition was modest, but in G2 ctc1 atr, telomeres shortened precipitously, and this loss coincided with a dramatic decrease in telomerase activity in G2 atr mutants. Zeocin treatment also triggered a reduction in telomerase activity, suggesting that the prolonged absence of ATR leads to a hitherto unrecognized DNA damage response (DDR). Finally, our data indicate that ATR modulates DDR in CST mutants by limiting chromosome fusions and transcription of DNA repair genes and also by promoting programmed cell death in stem cells. We conclude that the absence of CST in Arabidopsis triggers a multifaceted ATR-dependent response to facilitate maintenance of critically shortened telomeres, and eliminate cells with severe telomere dysfunction.

---

*Reprinted with permission from Boltz K.A., Leehy K., Song X., Nelson A.D.L., and Shippen D.E. 2012. ATR cooperates with CTC1 and STN1 to maintain telomeres and genome integrity in *Arabidopsis*. *Molecular Biology of the Cell* **22**, 1838-1848.

Introduction

A critical function of telomeres is to differentiate natural chromosome ends from DNA damage. The protective cap that defines the chromosome terminus consists of telomere binding proteins that associate with the double-stranded region, the single stranded 3' G-rich extension (G-overhang), or that bridge these two domains. The best characterized telomere capping complexes are shelterin in vertebrates and CST (Cdc13/Stn1/Ten1) in budding yeast. The six member shelterin complex spans both the double- and single-strand regions of the telomere (Palm and de Lange, 2008). Within shelterin, TRF2 and POT1 play leading roles in chromosome end protection (van Steensel *et al.*, 1998; Baumann and Cech, 2001). The CST complex associates exclusively with the G-overhang (Lin and Zakian, 1996), forming a heterotrimeric complex with structural similarity to RPA (Gao *et al.*, 2007; Sun *et al.*, 2009) A null mutation in any CST component is lethal, while other alleles trigger massive degradation of the telomeric C-strand causing grossly extended G-overhangs (Nugent *et al.*, 1996; Grandin *et al.*, 1997; Grandin *et al.*, 2001). Deletion of either the Stn1 or Ten1 ortholog in fission yeast leads to catastrophic loss of telomeric DNA and end-toend chromosome fusions (Martín *et al.*, 2007).

CST has recently been discovered in plants and vertebrates (Song *et al.*, 2008; Miyake *et al.*, 2009; Surovtseva *et al.*, 2009). STN1 and TEN1 are sequence homologs of the budding and fission yeast proteins (Song *et al.*, 2008; Miyake *et al.*, 2009; Price *et al.*, 2010). The third member of the complex, CTC1 (Conserved Telomere maintenance Component 1), is not a sequence homolog of Cdc13, although it shares functional

similarities. Like Cdc13, CTC1 physically interacts with STN1 as well as lagging-strand replication machinery (Casteel *et al.*, 2009; Miyake *et al.*, 2009; Surovtseva *et al.*, 2009; Price *et al.*, 2010). In addition, CTC1 in complex with STN1 and TEN1 binds single stranded DNA, but in a sequence-independent manner (Miyake *et al.*, 2009).

Ctc1 or Stn1 knockdown in human cells results in an increase in G-overhang signal, sporadic loss of telomeric DNA and aberrant chromatin bridges (Miyake *et al.*, 2009; Surovtseva *et al.*, 2009). Recent studies reveal that mutations in *CTC1* underly the rare human genetic disorder Coats plus, characterized by neurological and gastrointestinal defects (Anderson *et al.*, 2012). Coats plus patients also exhibit shortened telomeres and evidence of an ongoing DNA damage response (Anderson *et al.*, 2012). The major function for vertebrate CST may be related to DNA replication and repair, and not to chromosome end protection per se (Linger and Price, 2009; Giraud-Panis *et al.*, 2010; Price *et al.*, 2010). Recent studies show that *Xenopus* CST is required to prime ssDNA for replication (Nakaoka *et al.*, 2011). In addition, genetic data argue that CST and shelterin act in distinct pathways to promote telomere integrity in human cells. When both Stn1 and Pot1 are depleted, a synergistic increase in telomere dysfunction-induced foci is observed (Miyake *et al.*, 2009).

CST plays a pivotal role in protecting plant telomeres. Although *ctc1* and *stn1* null mutants are viable, they suffer dramatic telomere shortening, end-to-end chromosome fusions, increased G-overhangs and elevated extra-chromosomal telomeric circles, indicative of aberrant telomere recombination (Song *et al.*, 2008; Surovtseva *et al.*, 2009). Genetic analysis of *Arabidopsis thaliana STN1* and *CTC1*

confirms that these two components act in the same pathway for chromosome end protection (Surovtseva *et al.*, 2009). Unlike vertebrates, *Arabidopsis* harbors only a subset of shelterin components and thus far, none of these are required for chromosome end protection (Watson and Riha, 2010). Moreover, *Arabidopsis* encodes three POT1-like proteins, which associate with telomerase instead of the telomere (Surovtseva *et al.*, 2007; Cifuentes-Rojas *et al.*, 2011). Thus, CST appears to function as the major telomere protection complex in plants (Price *et al.*, 2010). CST is also likely to play a role in DNA replication in *Arabidopsis*, given its interaction with DNA polymerase α (Price *et al.*, 2010) and the results of vertebrate studies described above.

When telomere integrity is compromised due to loss of essential capping proteins, or prolonged inactivation of telomerase, the unprotected chromosome terminus triggers a cellular DNA damage response (DDR) that is mediated by the phosphoinositide-3-kinase-related protein kinases, ATM (Ataxia-Telangeictasia Mutated) or ATR (ATM and Rad3-related) (Sabourin and Zakian, 2008). ATM primarily responds to double-strand breaks, while ATR is activated by excessive single-stranded DNA (Nam and Cortez, 2011). As expected for telomere duplex binding components, TRF2 in vertebrates suppresses activation of ATM (Denchi and de Lange, 2007), while the single-strand binding proteins, mouse Pot1a (Denchi and de Lange, 2007), chicken Pot1 (Churikov *et al.*, 2006), and yeast Cdc13 (Garvik *et al.*, 1995; Ijpma and Greider, 2003; Hirano and Sugimoto, 2007), suppress an ATR-dependent DDR.

ATR and ATM are also required to maintain normal telomeres. Neither ATM nor ATR have been shown to affect telomerase enzyme activity levels in yeast or

vertebrates (Sprung *et al.*, 1997; Chan *et al.*, 2001; McNees *et al.*, 2010), but in yeast both kinases are implicated in the recruitment of telomerase to chromosome ends. In *Schizosaccharomyces pombe*, Tel1 (ATM) and Rad3 (ATR) are required for Ccq1-mediated interaction with telomerase (Moser *et al.*, 2009; Moser *et al.*, 2011). Similarly, in budding yeast Mec1 (ATR) and Tel1 (ATM) are each proposed to phosphorylate Cdc13 as a prerequisite for telomerase recruitment (Tseng *et al.*, 2006), although this finding is now controversial (Gao *et al.*, 2010). Nevertheless, a number of studies show that Tel1 facilitates the preferential recruitment of telomerase to critically shortened telomeres (Arneric and Lingner, 2007; Bianchi and Shore, 2007; Sabourin *et al.*, 2007), and stimulates telomerase repeat addition processivity on these chromosome ends (Chang *et al.*, 2007). Analysis of the ATR-deficient Seckel mouse indicates that while ATR is not required for telomerase recruitment to short telomeres (McNees *et al.*, 2010), it suppresses telomere fusions and the formation of fragile sites triggered by replication fork stalling in highly repetitive telomere repeat arrays (Martínez *et al.*, 2009; Sfeir *et al.*,2009; McNees *et al.*, 2010).

Many key components of DDR are conserved in plants, but there is considerable divergence in cell cycle regulated responses relative to vertebrates (Dissmeyer *et al.*, 2009). For example, ATM and ATR null mutations are not lethal in plants (Garcia *et al.*, 2003; Culligan *et al.*, 2004), and there is substantial overlap in the two pathways (Culligan *et al.*, 2004; Friesner *et al.*, 2005; Furukawa *et al.*, 2010). Moreover, plants are extraordinarily tolerant to genome instability, an outcome that may reflect the presence of undifferentiated stem cell niches in the shoot and root apical meristems. Meristematic

cells allow for continual growth and tissue differentiation, blunting the effect of DNA damage in somatic tissue. Ionizing radiation, for instance, will induce cell cycle arrest in meristems, but not in somatic cells (Hefner *et al.*, 2006).

Although mutation of either ATM or ATR has no effect on telomere length homeostasis in *Arabidopsis* (Vespa *et al.*, 2005), these kinases act synergistically with telomerase to maintain the telomere tract (Vespa *et al.*, 2005; Vespa *et al.*, 2007). Plants doubly deficient in ATM and TERT, the telomerase catalytic subunit, experience an abrupt, early onset of genome instability compared to *tert* single mutants (Vespa *et al.*, 2005). Analysis of individual telomere tracts showed that that ATM prevents stochastic deletional recombination events, allowing cells to maintain similar telomere lengths on homologous chromosome arms (Vespa *et al.*, 2007). ATR makes a more immediate contribution to telomere maintenance than ATM (Vespa *et al.*, 2005). From the outset, telomeres in double *atr tert* mutants shorten at a greatly accelerated pace relative to *tert*, so that telomere dysfunction occurs in the third generation of the double mutant, compared to the sixth generation of *tert*.

Here we employ a genetic approach to investigate how CST components interface with ATM and ATR to promote telomere integrity and genome stability in *Arabidopsis*. We demonstrate a pivotal role for ATR in the response to CST abrogation that leads to programmed stem cell death. We also show that the combined absence of ATR and CST results in catastrophic loss of telomere tracts in a biphasic manner. The second, more severe phase of telomere shortening coincides with strong down-regulation of telomerase activity. These findings indicate that ATR and CST act

synergistically to maintain genome integrity and telomere length homeostasis.

Materials and Methods

Plant Lines and Growth Conditions

Mutant *Arabidopsis thaliana* lines and genotyping have been previously described. The alleles used were *ctc1-1* and *ctc1-3* (Surovtseva *et al.*, 2009), *stn1-1* (Song *et al.*, 2008), *atr-2* (Culligan *et al.*, 2004), and *atm-2* (Garcia *et al.*, 2003). Crosses were made with plants heterozygous for *ctc1* or *stn1* and homozygous mutant for *atr* or *atm*. F1 plants were genotyped to identify plants that were heterozygous for both alleles. These were self-crossed and F2 siblings were used for analysis. Plants were grown on soil at 22°C under 16 h light/8 h dark conditions. For experiments using seedlings, seeds were sterilized in 50% bleach with 0.1% Triton-X 100 and then plated on MS with 0.7% agar (Caisson Labs). Plates were placed in the dark at 4°C for 2-4 days and then moved to long day conditions. For zeocin treatment, seeds were treated as described above. When seedlings were 5-7 days old, they were transferred to liquid MS culture either with or without 20μM zeocin (Invitrogen). Seedlings were grown in the dark for three days and then harvested for protein extraction.

Quantitative RT-PCR

Total RNA was extracted from G1 flowers using the E.Z.N.A. Plant RNA kit with on-column DNaseI digestion (Omega Bio-tek). To make cDNA, 2μg of RNA was used with the qScript cDNA Supermix (Quanta Biosciences). cDNA was diluted 1:4 in 10μg/ml yeast tRNA (Sigma) and 1μl was used in each qPCR reaction. The SsoFast EvaGreen Supermix (Bio-Rad) was used following manufacturers recommendations.

Reactions were run on a Bio-Rad CFX96 thermalcycler using 58°C primer annealing and 10s extension. RNA from at least three individual plants was used for each genotype and two replicates were run for each reaction. The raw amplification data was imported into LinRegPCR (Ruijter *et al.*, 2009) using the default settings. The window of-linearity and Cq threshold were calculated for each amplicon group. The resulting Cq values, which had been adjusted for the mean PCR efficiency for each amplicon, were used for calculation of expression levels. For each run, we measured three reference genes (*GAPDH, TIP41L*, and *At4G26410*) reported by Czechowski *et al*. (Czechowski *et al.*, 2005). The geometric mean of the three reference genes was used to calculate expression levels by the ΔΔCt method. Expression levels for each genotype were averaged and compared to wild type. Primers sequences were 5'-TGCATCCATTAAGTTGCCCTGTG-3' and 5'-TAGGCTGAGAGTGCAGTGGTTC-3' for *BRCA1* (At4G21070), 5'-ATGCTACTCTGGCACGGTTCAC-3' and 5'-AGGAGGAGCTATTCGCAGACCTTG-3' for *PARP1* (At4G02390), and 5'- CGAGGAAGGATCTCTTGCAG-3' and 5'-GCACTAGTGAACCCCAGAGG-3' for *RAD51* (At5G20850).

Telomere length measurement, in-gel hybridization, TF-PCR and TRAP

Genomic DNA was extracted from whole plants or seedlings using 2x CTAB buffer (Vespa *et al.*, 2005) with slight modification. Plant extracts were incubated for 1 h at 50°C, and all mixing was done by inverting tubes rather than vortexing. TF- PCR and PETRA (Heacock *et al.*, 2004) and TRF (Fitzgerald *et al.*, 1999) were conducted as

previously reported. For all three assays, products were detected by Southern Blot with a [32P]-5'-end-labeled (TTTAGGG)4 probe. A [32P]-5'-end-labeled (CCCTAAA)3 probe was used for in-gel hybridization as described previously (Surovtseva *et al.*, 2009). Telomere lengths from PETRA analyses were calculated using QuantityOne software (Bio-Rad). For lanes with multiple bands, the average size was calculated. Protein extracts from 5 to 7 day-old seedlings were used for quantitative TRAP as previously described (Kannan *et al.*, 2008).

Propidium iodide staining and cytogenetics

Five to seven day-old G2 seedlings were gently removed from MS plates and placed in 10 μg/ml propidium iodide solution diluted in water for 10 min at room temperature in the dark. Seedlings were then transferred to water. Roots and shoots were separated and roots were mounted on slides in water. *Arabidopsis* chromosome spreads were prepared from pistils as described (Riha *et al.*, 2001). The spreads were mounted on slides with Vectashield Plus DAPI (Vector Laboratories). All slides were visualized with a Zeiss Axioplan2 epifluorescent microscope using a rhodamine filter for PI slides and a DAPI filter for chromosome spreads. ImageJ was used to adjust the brightness and contrast of images.

Results

Loss of ATR rescues morphological defects in CST mutants.

To explore the role of ATR and ATM in plants lacking CST, we crossed *ctc1* or *stn1* heterozygotes to *atr* and *atm* mutants. F1 plants heterozygous for both mutations were self-crossed and offspring were used for analysis. As previously reported (Garcia *et al.*, 2003; Culligan *et al.*, 2004; Vespa *et al.*, 2005), *atm* (Figure 1A) and *atr* (Figure 1B) mutants were phenotypically indistinguishable from wild type. In contrast, *ctc1* and *stn1* mutants exhibited serious morphological defects (Song *et al.*, 2008; Surovtseva *et al.*, 2009), including fasciated inflorescence bolts and flowers (Figure 1C, arrowheads; Supplemental Figure 1, white arrows), irregularly spaced siliques (Figure 1C, arrows; Supplemental Figure 1), and small curved leaves. Although *ctc1* and *stn1* mutants always display morphological abnormalities, the expressivity of the mutant alleles is somewhat variable, with some individuals showing more severe phenotypes than others (Song *et al.*, 2008; Surovtseva *et al.*, 2009). Both *ctc1 atm* and *stn1 atm* double mutants displayed the same range of growth defects as *ctc1* (Figure 1A) or *stn1* mutants (Supplemental Figure 1A). In contrast, *ctc1 atr* and *stn1 atr* mutants showed only minor perturbations in morphology, mainly irregularly spaced siliques. Approximately 30% of the double mutants appeared wild type (Figure 1B and C, Supplementary Figure 1B). The apparent rescue of morphological defects in *ctc1 atr* and *stn1 atr* mutants is consistent with the conclusion that CST protects against ATR activation.

The improvement of morphological deficiencies in *ctc1 atr* mutants was only

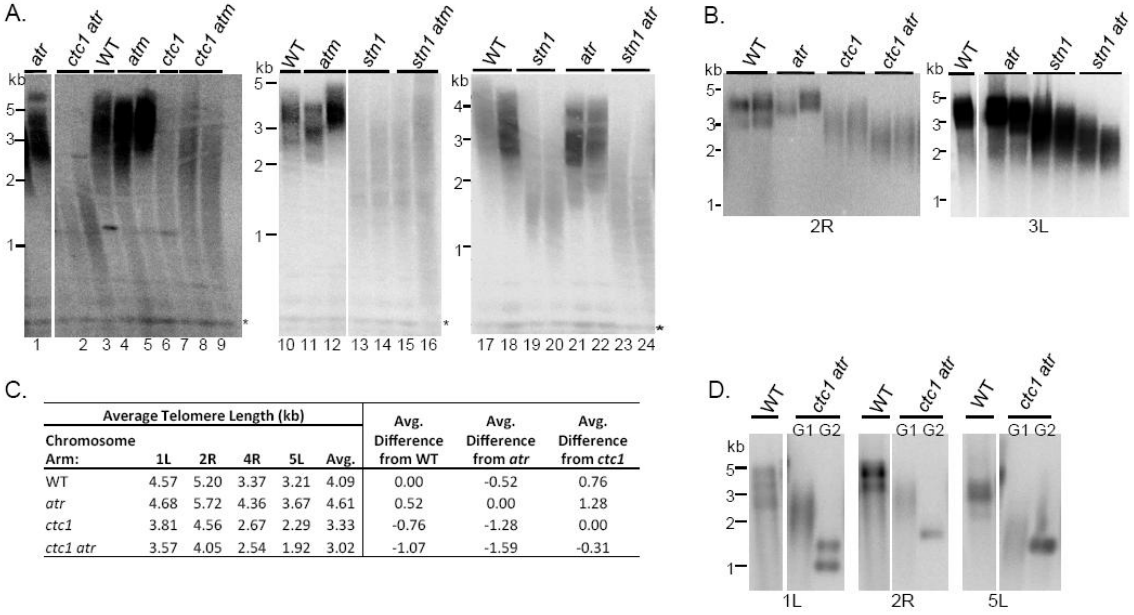
temporary. Second generation (G2) *ctc1 atr* mutants showed severe developmental defects, and most died before bolting (Figure 1D and E). Many of the phenotypes associated with G2 *ctc1 atr* resembled G1 *ctc1* mutants (Surovtseva *et al.*, 2009). Defects included curved, misformed leaves and severe floral abnormalities, such as missing anthers, curved pistils, open carpels with seeds exposed and petals that were green like sepals (Figure 1E). We conclude that ATR alters plant growth in response to CST abrogation.

Figure A3-1. Loss of ATR rescues the morphological defects of *ctc1* mutants. The morphology of *ctc1* mutants in the presence or absence of ATM or ATR is shown. (A) The phenotype of a *ctc1 atm* double mutant (right) resembles the *ctc1* single mutant. (B and C) Morphological defects of *ctc1* mutants are largely rescued when ATR is lost. Arrowheads indicate fasciated stems and flowers; arrows indicate irregular phyllotaxy. Images of second generation (G2) *ctc1 atr* mutants are presented showing an intact plant (D) with curved, small leaves, or malformed flowers (E) bearing a curved pistil, and stamen and petal deficiency.

ATR facilitates telomere length maintenance in the absence of CTC1 or STN1.

The morphological rescue seen in CST mutants lacking ATR argues that ATR is activated by telomere dysfunction. Given the role of ATR in telomere maintenance in telomerase mutants (Vespa *et al.*, 2005), we considered the possibility that ATR also contributes to telomere maintenance in plants lacking CST. Bulk telomere length was monitored using Terminal Restriction Fragment (TRF) analysis. As previously reported (Vespa *et al.*, 2005), telomere tracts in *atr* and *atm* were similar to wild type (Figure 2A, lanes 1, 4, 6), while G1 *ctc1* telomeres were shorter and more heterogeneous (Figure 2A, lane 7). The absence of ATM did not affect telomere length in G1 *ctc1* mutants (Figure 2A, lanes 8-9). In both G1 *ctc1* and G1 *ctc1 atm* mutants, telomeres ranged from 1-5kb, with a peak signal at 2kb. In contrast, telomeres were consistently shorter in G1 *ctc1 atr* mutants than in G1 *ctc1* (Figure 2A, lanes 2-3 and 7), with some signals trailing below 1kb (peak = 1.5kb). Similar findings were obtained with G1 *stn1 atm* (Figure 2A, lanes 13-16) and G1 *stn1 atr* mutants (Figure 2A, lanes 19-20 and 23-24).

Primer Extension Telomere Repeat Amplification (PETRA) was employed to precisely measure telomere length on individual chromosome arms. In this assay, wild type telomeres range from 2-5kb and typically appear as one to three bands depending on the chromosome arm (Figure 2B) (Heacock *et al.*, 2004). As with bulk telomere analysis, PETRA showed that the telomere profiles of *atr* (Figure 2B) and *atm* (Supplemental Figure 2A and B) were similar to wild type, whereas telomeres from G1 *ctc1* and G1 *stn1* migrated as a broad smear ranging from 1.5-4kb (Figure 2B). PETRA confirmed that telomere tracts were similar in G1 *ctc1* and G1 *ctc1 atm* mutants



Chromosome Arm:	Average Telomere Length (kb) 1L	2R	4R	5L	Avg.	Avg. Difference from WT	Avg. Difference from *atr*	Avg. Difference from *ctc1*
WT	4.57	5.20	3.37	3.21	4.09	0.00	-0.52	0.76
atr	4.68	5.72	4.36	3.67	4.61	0.52	0.00	1.28
ctc1	3.81	4.56	2.67	2.29	3.33	-0.76	-1.28	0.00
ctc1 atr	3.57	4.05	2.54	1.92	3.02	-1.07	-1.59	-0.31

Figure A3-2. ATR, but not ATM, contributes to telomere length maintenance in *ctc1* and *stn1* mutants. (A) TRF analysis of *ctc1* crosses to *atr* and *atm* (lanes 1-9) and *stn1* crosses to *atm* (lanes 10-16) and *atr* (lanes 17-24). (B) PETRA results for the 2R telomere in *ctc1 atr* mutants and the 3L telomere in *stn1 atr* mutants. (C) Quantification of telomere lengths from *ctc1 atr* PETRA analysis shown in panel B. Telomere length was calculated by subtracting the distance of the subtelomeric primer binding site relative to start of the telomere repeat array from the PETRA value. For all genotypes, n=4. (D) Parent-progeny PETRA analysis of telomeres in G1 and G2 *ctc1 atr* mutants. Asterisk indicates interstitial telomeric repeats used as a loading control.

(Supplemental Figure 2A). In contrast, telomeres in G1 *ctc1 atr* mutants were shorter by an average of 300 bp compared to G1 *ctc1* mutants (Figure 2B and 2C). The same result was obtained for *stn1* mutants in both *atm* (Supplemental Figure 2B) and *atr* (Figure 2B) deficient backgrounds. Hence, ATR, but not ATM, contributes to telomere length maintenance when CST is compromised.

We examined the status of the G-overhang in G1 *ctc1 atr* mutants using in-gel hybridization. This assay detects single-stranded G-rich telomeric DNA either at the extreme chromosome terminus or within the double-stranded telomere region, if gaps are present in the C-strand. As previously reported (Surovtseva *et al.*, 2009), *ctc1* single mutants showed enhanced G-overhang signals, three- to six- fold greater than wild type (Supplemental Figure 3). G-overhang status was wild type in *atr* mutants. Furthermore, the loss of ATR did not exacerbate the G-overhang phenotype in *ctc1* mutants (Supplemental Figure 3). We conclude that ATR does not play a significant role in G-overhang maintenance, and further that *ctc1 atr* mutants do not carry extensive sections of incompletely replicated telomeric C-strand DNA.

Since G2 *ctc1 atr* mutants have much more severe morphological defects than G1 *ctc1 atr* (Figure 1D and E), we were prompted to examine telomere length in G2 double mutants using PETRA. Telomere tracts in G2 *ctc1 atr* were much shorter (up to 1kb) than their G1 parents (Figure 2D). This attrition is more than three times greater than the telomere shortening in G1 *ctc1 atr* mutants versus their *ctc1* siblings (300 bp) (Figure 2, A-C), and more than two times higher than G2 *stn1* mutants versus their G1 parent (~400 bp)(data not shown). In conjunction with telomere shortening, the profile

of telomere fragments switched from heterogeneous, smeary bands in the G1 *ctc1 atr* parents to very homogenous, sharp bands in the G2 *ctc1 atr* offspring (Figure 2D). PETRA assays conducted with five generations of *atr* mutants revealed no change in telomere length (Supplemental Figure 2C), confirming that the telomere maintenance defect in *ctc1 atr* mutants reflects a synergistic effect of both ATR and CST dysfunction. These data further indicate that ATR contributes to telomere maintenance in a biphasic manner. In the first generation of a CST deficiency, ATR makes a modest contribution to telomere maintenance. However, the prolonged absence of ATR in plants lacking CST leads to a much more dramatic loss of telomeric DNA.

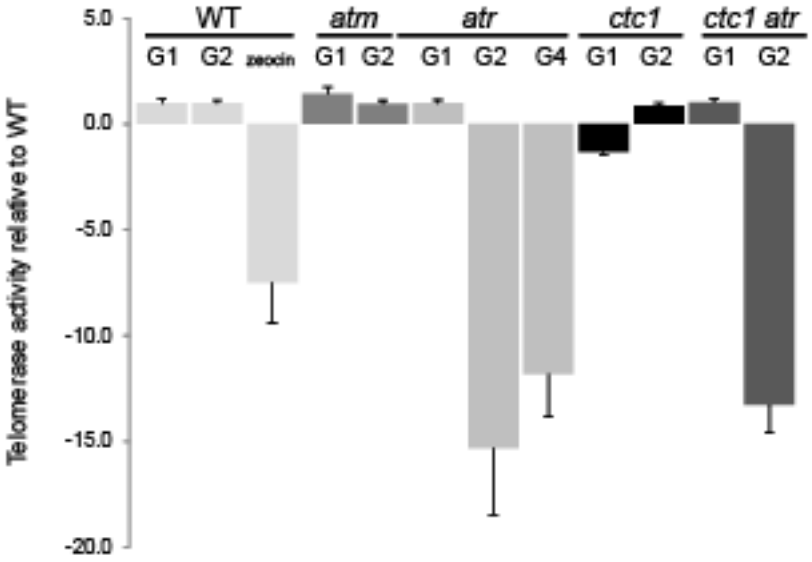


Figure A3-3. ATR stimulates telomerase activity. Quantitative TRAP results for first (G1), second (G2) and fourth (G4) generation mutants of different genotypes are shown. Q-TRAP was also performed on wild type seedlings treated with 20µM zeocin for 3 days. All samples were from flowers except G2 *atr,* G2 *ctc1*, and G2 *ctc1 atr*, which were from seedlings. Telomerase activity is plotted relative to wild type. For zeocin-treated seedlings, telomerase activity is relative to untreated wild type seedlings. Error bars represent standard deviation. n=2 for all genotypes except G1 WT n=5, zeocin-treated WT n=6, G1 *ctc1* n=4, G2 *atr* n=3, and G4 *atr* n=4.

Inactivation of ATR down-regulates telomerase enzyme activity.

A profile of shorter, more homogeneous telomere tracts is consistent with a defect in telomerase-mediated telomere maintenance (Riha *et al.*, 2001; Kannan *et al.*, 2008). Thus, one explanation for the enhanced rate of telomere loss in G2 *ctc1 atr* mutants is that telomerase can no longer act on dysfunctional chromosome ends. To investigate this possibility, we used Quantitative Telomere Repeat Amplification Protocol (Q-TRAP) to measure telomerase enzyme activity levels in consecutive generations of *ctc1 atr* mutants. As expected (Song *et al.*, 2008; Surovtseva *et al.*, 2009), telomerase activity was robust in G1 and G2 *ctc1* and *stn1* seedlings, and indistinguishable from wild type samples (Figure 3). Wild type levels of telomerase activity were also detected in G1 *atr* mutants. Unexpectedly, however, telomerase activity declined by ~15-fold in G2 *atr* mutants (Figure 3). This decrease persisted in subsequent plant generations with G4 *atr* mutants also exhibiting dramatically reduced enzyme activity. The reduction in telomerase activity was not confined to a specific developmental stage; Q-TRAP data obtained from both seedlings and flowers gave similar results (Figure 3). Notably, Q-TRAP revealed the same level of enzyme activity in G1 *ctc1 atr* mutants as in wild type plants, and enzyme activity in G2 *ctc1 atr* decreased by the same amount as in G2 *atr* (Figure 3). Hence, loss of ATR, and not CTC1, leads to decreased telomerase activity.

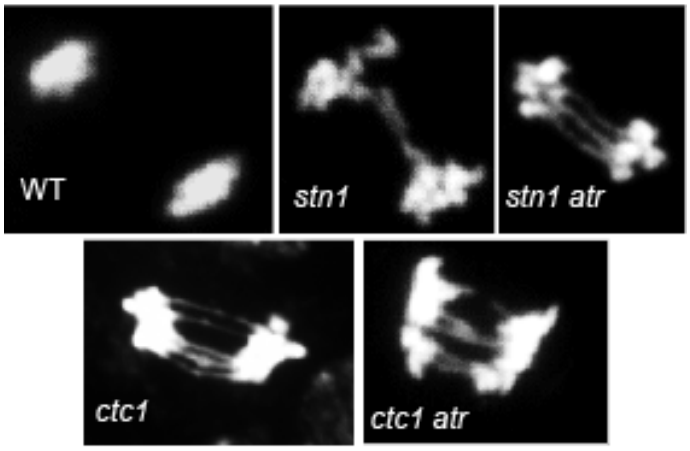
In yeast and vertebrates, disruption of ATR causes genome wide replicative stress (Nam and Cortez, 2011), suggesting that the stimulus for reduced telomerase activity in G2 *atr* mutants might be accumulating genome damage. To investigate

whether genotoxic stress triggers a decrease in telomerase activity, wild type seedlings were treated with zeocin, which induces double-strand breaks. Q-TRAP revealed ~7.5-fold reduction in telomerase in treated seedlings versus controls (Figure 3). This observation suggests that the repression of telomerase activity in G2 *atr* mutants may reflect the activation of a DDR triggered by replicative stress. Altogether, these results show that the dramatic loss of telomeric DNA in G2 *ctc1 atr* mutants correlates with an abrupt decline in telomerase enzyme activity.

ATR suppresses the formation of end-to-end chromosome fusions in CST mutants.

Catastrophic loss of telomeric DNA in *ctc1* and *stn1* mutants coincides with the onset of telomere fusions (Song *et al.*, 2008; Surovtseva *et al.*, 2009). Dysfunctional telomeres are recruited into chromosome fusions through the non-homologous end-joining (NHEJ) pathway, which is activated by ATM and indirectly by ATR (Denchi and de Lange, 2007; Deng *et al.*, 2009). Therefore, we asked if the accelerated telomere shortening in plants lacking CST and ATR correlates with an increased incidence of telomere fusions using telomere fusion PCR (TF-PCR). TF-PCR employs primers specific to unique subtelomeric sequences on each chromosome arm to amplify junctions of covalently fused telomeres. For these studies, DNA from mature G1 mutants was analyzed. As expected, telomere fusions were not observed in wild type, *atr* (Supplemental Figure 4B, D) or *atm* (Supplemental Figure 4A, C) mutants. In contrast, massive chromosome end-joining events, represented by abundant heterogeneous smears, were associated with the loss of CTC1 (Supplemental Figure 4A-B) or STN1 (Supplemental Figure 4C-D). When either ATR (Supplemental Figure

A.



B.

	Anaphases		
Genotype	Total	Bridges	Ratio %
WT	69	0	0%
stn1	369	76	21%
ctc1	166	38	23%
atm	202	0	0%
atr	454	6	1%
stn1 atm	137	27	20%
stn1 atr	501	288	57%
ctc1 atr	234	125	53%
stn1 atm atr	323	225	70%

Figure A4-4. End-to-end chromosome fusions increase in plants lacking CST and ATR. (A) Cytology of anaphases from pistils from G1 plants of the genotypes indicated. Spreads are stained with DAPI. (B) Quantification of anaphase bridges from cytology in (A).

4B, D) or ATM (Supplemental Figure 4A, C) was absent in *ctc1* or *stn1* mutants, TF-PCR products were still detected. TF-PCR provides an indication of whether telomeres are prone to end-joining reactions, but it does not give quantitative information about the number of chromosome fusions. To obtain a quantitative assessment of telomere joining events, we monitored the incidence of anaphase bridges in mitotically dividing cells using conventional cytology (Figure 4A). As described previously (Song *et al.*, 2008; Surovtseva *et al.*, 2009), bridged chromosomes were detected in the floral pistils of G1 *ctc1* and *stn1* mutants (23% and 21% of all anaphases, respectively), compared to few or none in wild-type, *atr*, and *atm* mutants (Figure 4B). The loss of ATM did not alter the percentage of anaphase bridges in *stn1* mutants. Conversely, there was a dramatic increase in the incidence of anaphase bridges in G1 *stn1 atr* (57%) and G1 *ctc1 atr* (53%) relative to *stn1* and *ctc1* (Figure 4B). Remarkably, 70% of anaphases in the triple G1 *stn1 atr atm* mutants contained bridged chromosomes (Figure 4B). Thus, an ATR- and ATM-independent mechanism can promote fusion of dysfunctional telomeres. The increased incidence of chromosome bridges suggests that ATR inhibits telomere fusion in CST mutants.

ATR attenuates the transcriptional response to DNA damage in plants lacking CTC1.

The role of ATR in repressing telomere fusions together with the accelerated telomere shortening, and morphological disruptions in CST mutants argues that loss of CST triggers an ATR-mediated DDR. To investigate this possibility, we monitored the expression of several transcripts implicated in DDR (*RAD51*, *BRCA1* (*BREAST*

CANCER SUSCEPTIBILITY 1) and *PARP1* (*Poly [ADP-ribose] polymerase 1*)) (Doucet-Chabeaud *et al.*, 2001; Lafarge and Montané, 2003; Yoshiyama *et al.*, 2009). Quantitative RT-PCR was performed using cDNA made from first generation (G1) *ctc1* flowers. Expression of both *PARP1* and *BRCA1* was significantly up-regulated in *ctc1* mutants compared to wild type (3.7- and 1.9-fold, respectively) (Figure 5). In addition, *RAD51* expression was 1.5 times higher in *ctc1* mutants (Figure 5), but the difference was not statistically significant. These results suggest that the CST complex protects against a DDR.

We next asked if ATM or ATR are necessary to initiate a transcriptional response in plants lacking CST since in *Arabidopsis*, the response to double-strand breaks is mostly mediated by ATM, but ATR is also required (Friesner *et al.*, 2005). In *ctc1 atm* mutants, *PARP1* and *BRCA1* transcripts were above wild type levels (2.1 and 1.7 times wild type, respectively), but were slightly less abundant than in *ctc1* mutants. This finding suggests that ATM contributes to the activation of a DNA repair transcriptional program in *ctc1* mutants. A more dramatic change in transcript level was observed in plants doubly deficient in CTC1 and ATR. Expression of all three DDR genes was significantly elevated in *ctc1 atr* mutants relative to wild type, *atr* or *ctc1* (Figure 5). Compared to wild type, *ctc1 atr* mutants showed a 7.7 fold increase in *PARP1* expression, a 2.3-fold increase in RAD51, and a 3.1-fold increase in BRCA1. Thus, ATR curbs the transcriptional response to loss of CTC1. This observation is consistent with ATR-mediated suppression of chromosome fusions.

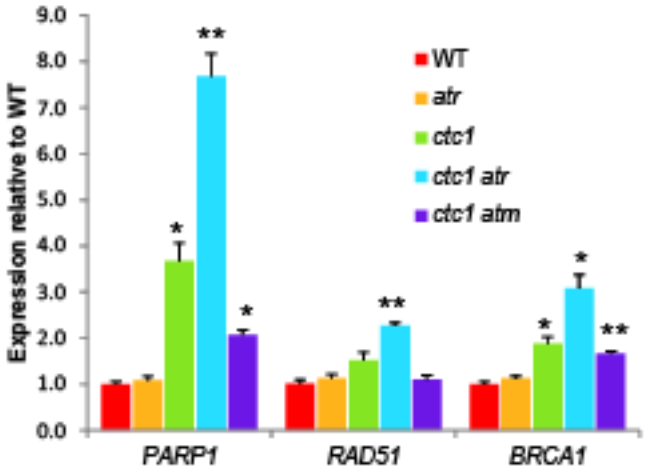


Figure A3-5. Loss of CTC1 activates an ATR-dependent transcriptional response. qRT-PCR results are shown for the DDR transcripts *PARP1*, *BRCA1*, and *RAD51* in floral organs. Expression levels are relative to wild type, and data for first generation (G1) mutants are shown. For each genotype, n=3, except for *ctc1 atm*, n=2. Single asterisk denotes a p-value <0.05 relative to wildtype; two asterisks denote a p-value<0.005 relative to wildtype (Student's T-test). Error bars represent S.E.M.

ATR promotes programmed cell death in ctc1 mutants.

ATR is implicated in programmed cell death signaling in *Arabidopsis* (Fulcher and Sablowski, 2009; Furukawa *et al.*, 2010). To further explore the role of ATR in plants lacking CST, we monitored stem cell viability in root apical meristems (RAM) of seedlings using propidium iodide (PI) staining (Figure 6A). PI is a membrane-impermeable dye that is excluded from live cells. In dead cells, PI passes through the cell membrane and binds nucleic acids. The limited biomass of young seedlings precluded genotyping to identify G1 double mutants so early in their development. Therefore, we examined the RAM in their progeny, G2 *ctc1 atr* mutants. As expected, PI staining was not associated with the RAM in wild type seedlings (Figure 6A, panel ii). Similarly, G2 *atr* seedlings showed no PI staining (Figure 6A, panel iii). In contrast, strong PI staining was observed in G2 *ctc1* RAM (Figure 6A, panel iv) or G2 *stn1* RAM (Figure 6A, panel v), consistent with activation of a robust DDR. We next asked if ATR is responsible for cell death in CST mutants (Figure 6A, panel vi). Strikingly, the number of PI positive cells in G2 *ctc1 atr* dropped to an average 1.75 cells/root compared to 5.75 and 4.35 cells/root for *stn1* and *ctc1*, respectively (Figure 6A, panel vi and Figure 6B). A subset of mutant seedlings (25% in *stn1*, 35% in *ctc1*, and 67% in *ctc1 atr*) had no PI-positive cells. The short roots from these plants had a high density of root hairs and no obvious RAM (Figure 6A, panels vii-viii). We speculate that in such plants, epithelial precursor cells may be able to differentiate, but other cell types have been eliminated from the RAM or have differentiated inappropriately. These mutant roots are remarkably similar to gamma-irradiated *lig4* roots, where RAM cells are

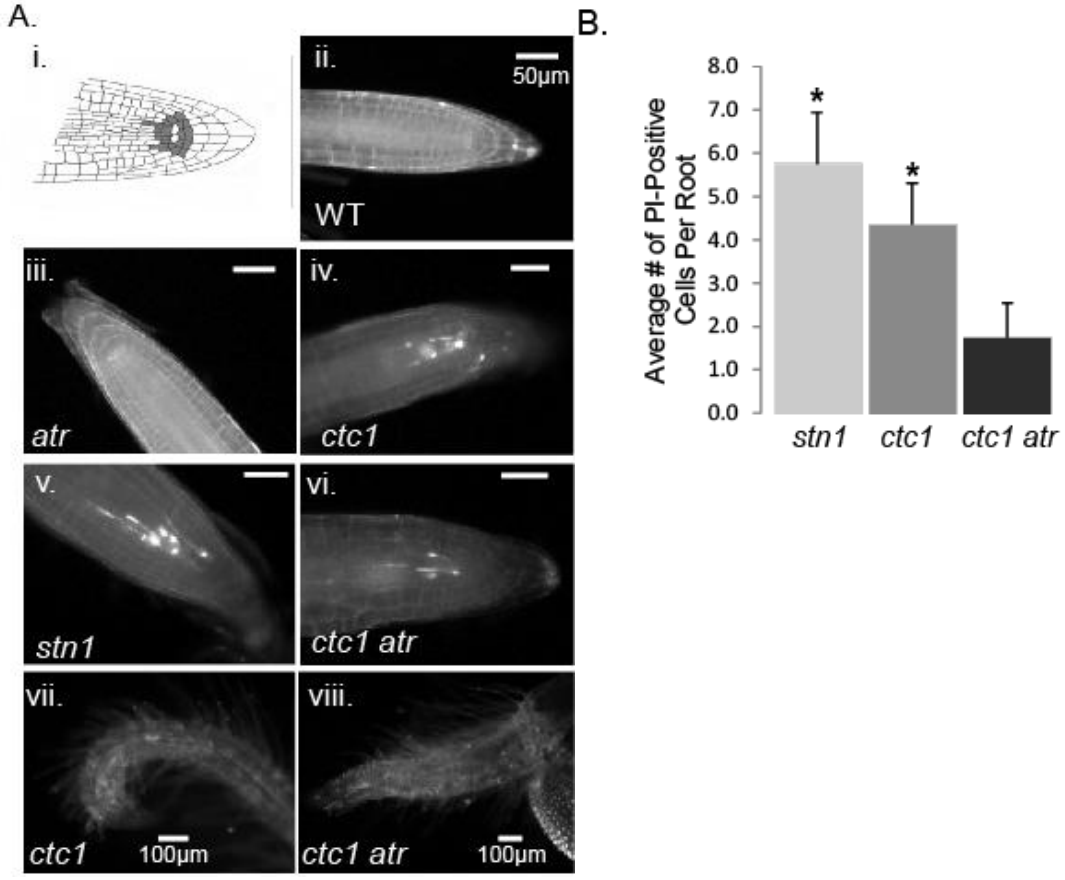


Figure A3-6. ATR activates programmed cell death of the root apical meristem (RAM) of *ctc1* mutants. (A) Representative images of G2 seedling root tips stained with propidium iodide (PI). (i) Diagram of a root tip. Stem cells and adjacent daughter cells are shaded gray. White cells in the RAM center are quiescent center cells. WT (ii) and *atr* (iii) roots are PI-negative, but the RAM of *ctc1* (iv) and *stn1* (v) mutants have numerous PI-positive (dead) cells. (vi) Fewer PI-positive cells are present in *ctc1 atr* mutants. (vi-vii) A subset of *ctc1* or *stn1* roots were PI-negative, but displayed severe morphological defects. (B) Quantification of PI-positive cells in different genetic backgrounds. The average number of PI-positive cells per root tip is shown. *stn1* (n= 12); *ctc1* (n=17); *ctc1 atr* (n=12). Asterisk denotes a p-value <0.05 (Student's T-test). Error bars represent S.E.M.

arrested (Hefner *et al.*, 2006). Taken together, these data indicate that ATR activationleads to programmed cell death in plants lacking CST. Further, we speculate that the decrease in PCD in *ctc1 atr* mutants leads to an accumulation of cells exhibiting DDR and increased numbers of end-to-end chromosome fusions.

Discussion

CST protects telomeres from activating ATR.

A key function of intact telomeres is to prevent the chromosome terminus from eliciting a cellular DDR that leads to end-to-end chromosome fusions and genome wide instability. Here we show that the *Arabidopsis* CST prohibits the activation of ATR-mediated DDR. We find that the absence of CTC1 results in elevated levels of DDR transcript expression and programmed cell death in the RAM. The sacrifice of stem cells by programmed cell death is a common response to DNA damage in plants (Fulcher and Sablowski, 2009; Furukawa *et al.*, 2010), and has obvious benefits for organismal viability. Several observations support the idea that ATR-mediated programmed cell death reduces genome instability in CST mutants. First, expression of DDR transcripts increases in *ctc1 atr* mutants compared to *ctc1* mutants. Second, the incidence of chromosome fusions increases in *ctc1 atr* mutants. Finally, plants lacking core components of CST display severe morphological abnormalities as a consequence of profound genome instability (Song *et al.*, 2008; Surovtseva *et al.*, 2009), and these phenotypes are largely rescued by a deficiency in ATR, but not ATM. The rescue is only temporary, however, and in the next generation (G2), *ctc1 atr* mutants suffer even more devastating developmental defects than G2 *ctc1* single mutants. This observation is

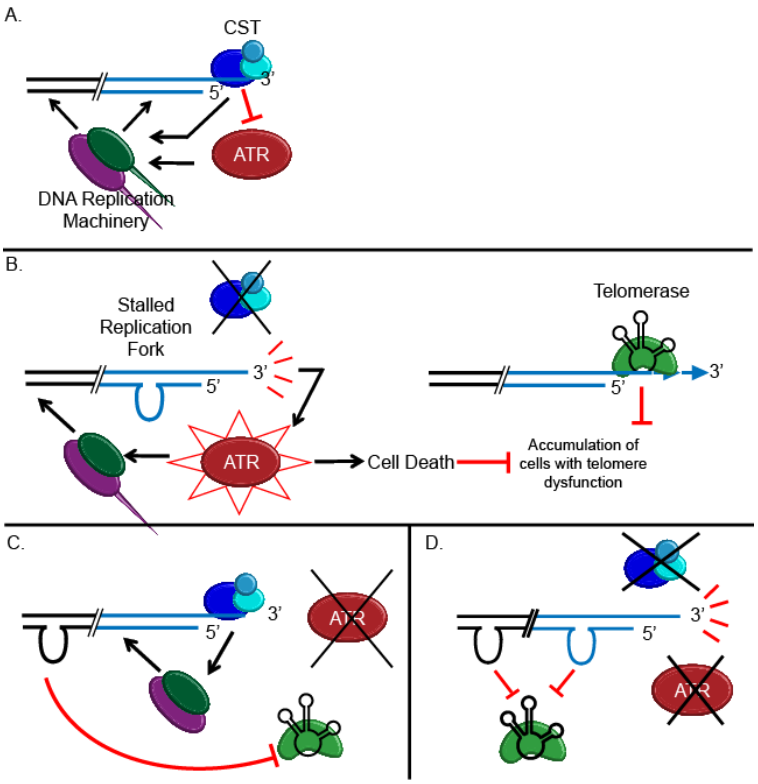


Figure A3-7. Model depicting CST and ATR cooperation in maintaining telomeric DNA and genome integrity in Arabidopsis. (A) In wild type plants, CST interacts with the 3' overhang to protect the chromosome terminus from telomere shortening, end-to-end chromosome fusions (Song et al., 2008; Surovtseva et al., 2009) and activation of ATR-dependent DDR (this study). ATR facilitates replication fork progression. Similarly, CST is thought to stimulate replication fork restart within the telomeric duplex via interaction with DNA polymerase alpha (Price et al., 2010; Nakaoka et al., 2011). Telomeric DNA is represented by blue lines. (B) Plants lacking CST activate ATR-dependent DDR, initiating programmed cell death in stem cell niches. Replication fork progression is perturbed in the telomeric duplex, contributing to the loss of telomeric DNA. Telomerase action delays the onset of complete telomere failure. (C) Accumulating replicative stress in atr mutants triggers an ATR-independent DDR that results in telomerase inhibition. Telomeres in the wild type size range can be maintained. (D) Catastrophic telomere shortening occurs in plants lacking both CST and ATR due incomplete replication of the duplex and failure of telomerase to act on critically shortened telomeres. See text for details.

consistent with checkpoint bypass, resulting in the accumulation of DNA damage when ATR is lost in *ctc1* mutants. We postulate that the failure to initiate programmed cell death allows *ctc1 atr* cells with dysfunctional telomeres to continue cycling until rampant genome instability leads to developmental arrest (Figure 7B).

While this manuscript was under review, Amiard *et al.* published a study that verifies and complements our findings concerning the role of CST in suppressing an ATR-mediated DDR (Amiard *et al.*, 2011). These authors show an ATR-dependent induction of γH2AX at telomeres in *Arabidopsis ctc1* mutants, consistent with our transcriptional data showing induction of DDR transcripts in response to loss of CTC1. Amiard and colleagues also demonstrate that ATR and ATM repress formation of anaphase bridges and promote PCD in *ctc1* mutants. They conclude, as do we, that ATR maintains genome stability in CST mutants (Amiard *et al.*, 2011).

Together, these *Arabidopsis* studies highlight the complexity of the DDR in plants and show that multiple, overlapping mechanisms are harnessed to detect and to process dysfunctional telomeres. For example, the increased incidence of telomere fusions in plants lacking CST and ATR could reflect survival of cells with profound telomere dysfunction due to checkpoint bypass, as well as a contribution of ATR in facilitating maintenance of short telomeres (see below). Notably, telomere fusions accumulate even in the absence of both ATM and ATR when CST is compromised (Amiard *et al.*, 2010; this study). A third PIKK family member in vertebrates, DNA-dependent protein kinase catalytic subunit (DNA-PKcs), functions in non-homologous end-joining (NHEJ) (Lieber *et al.*, 2003) and could potentially serve as back-up

mechanism to trigger telomere fusion. Plants lack an obvious DNA-PKcs ortholog, and thus the ATR/ATM independent response elicited by telomere dysfunction is unknown.

Further complicating matters, uncapped telomeres engage both canonical and noncanonical DNA repair pathways in *Arabidopsis*. Critically shortened telomeres fuse in the absence of two core NHEJ repair proteins, Ku70 and Ligase IV (Heacock *et al.*, 2007), and in plants lacking Ku as well as Mre11 (Heacock *et al.*, 2004). In humans, an alternative end-joining pathway, which employs PARP1 and DNA ligase III, is activated if the canonical DNA-PKcs/Ku pathway is non-functional (Audebert *et al.*, 2004). It is unknown if PARP1 plays a similar role in plants, but it is an intriguing possibility given the induction of *PARP1* expression in *ctc1* and *ctc1 atr* mutants (Figure 5).

Cooperation of CST and ATR in telomere maintenance

Figure 7 presents a model summarizing the multifunctional roles of ATR at *Arabidopsis* telomeres. The data presented here showing a central role for ATR in the response to CST abrogation provides additional support for the proposal that CST binds single-stranded DNA at the chromosome terminus in multicellular organisms (Miyake *et al.*, 2009; Surovtseva *et al.*, 2009) (Figure 7A). While our findings do not specifically address whether CST directly contacts the G-overhang, they are consistent with this conclusion and with the current model that single-strand telomere binding proteins protect the chromosome terminus by excluding RPA from the G-overhang (Gong and de Lange, 2010; Flynn *et al.*, 2011).

Our results show that CST and ATR cooperate in the maintenance of telomeric

DNA. We found that inactivation of ATR, but not ATM, accelerates the attrition of telomeric DNA at telomeres lacking CST. Multi-generational analysis of *ctc1 atr* mutants demonstrated that ATR makes a biphasic contribution to telomere length homeostasis. Our data indicate that in the first generation of a CST deficiency, the role of ATR is relatively minor. Telomeres are ~300bp shorter in *ctc1 atr* mutants than when ATR is intact. However, in the next generation, telomere shortening is much more aggressive, and up to 1kb more telomeric DNA is lost. We hypothesize that this biphasic response reflects two distinct contributions of ATR in promoting telomere maintenance (Figure 7B and 7C).

Emerging data indicate that ATR and CST cooperate to facilitate DNA replication through the telomeric duplex (Price *et al.*, 2010) (J. Stewart and C. Price, personal communication). ATR is activated in response to replication fork stalling (Verdun *et al.*,2005; Miller *et al.*, 2006), and specifically suppresses telomere fragility derived from incomplete replication (Martínez *et al.*, 2009; Sfeir *et al.*, 2009; McNees *et al.*, 2010). Notably, mammalian chromosomes depleted of CTC1 or STN1 display multiple telomere signals, consistent with telomere fragile sites (Price *et al.*, 2010). CST is proposed to participate in replication fork restart via its interaction with DNA polymerase-alpha (Casteel *et al.*, 2009; Price *et al.*, 2010). Consistent with this model, *Xenopus* CST is required for priming replication of ssDNA (Nakaoka *et al.*, 2011). Altogether these findings indicate CST and ATR cooperate in relieving replicative stress within the telomere duplex (Figure 7B and 7C). When both CST and ATR are compromised, replication fork stalling is increased (Figure 7D), triggering double-strand

breaks, and in turn, deletion of telomeric DNA.

Replicative stress may account for the modest increase in telomere shortening in G1 *ctc1 atr* mutants. While the extent to which ATR and CST modulate replication of the telomeric duplex in plants is unknown, preliminary data suggest that the contribution of these two components could be less significant in plants than in vertebrates. In human cells lacking CST, a small fraction of G-rich telomeric single-stranded DNA signal is resistant to exonuclease treatment (Surovtseva et al., 2009; Miyake *et al.*, 2009), consistent with incomplete replication of internal telomeric DNA tracts. Parallel analysis in Arabidopsis failed to detect exonuclease-resistant G-rich single-stranded DNA (Surovtseva et al. 2009), suggesting that CST acts primarily at the extreme chromosome terminus. We also found no increase in G-rich single-stranded DNA in *ctc1 atr* mutants relative to *ctc1*, implying that loss of ATR does not trigger massive replication fork stalling in CST mutants.

Telomerase and ATR

What accounts for the abrupt and dramatic loss of telomeric DNA in G2 *ctc1 atr* mutants? We propose that this delayed response reflects telomerase inhibition triggered by prolonged ATR inactivation. Depletion of ATR in mice leads to extensive chromosome fragmentation and a null mutation is embryonic lethal (Brown and Baltimore, 2000; de Klein *et al.*, 2000). In contrast, plants lacking ATR are viable, fully fertile and morphologically wild type (Culligan *et al.*, 2004). Although no overt genome instability is associated with ATR depletion in *Arabidopsis*, we speculate that accumulating replicative stress elicits a hitherto unrecognized DDR, one consequence

of which is telomerase repression (Figure 7C). In support of this hypothesis, we showed that the genotoxin zeocin inhibits telomerase activity in wild type seedlings. Strikingly, telomerase activity is unaffected in plants lacking CST, indicating that telomere dysfunction does not inhibit telomerase. Sustained repeat incorporation onto compromised chromosome ends would be advantageous if it delays the onset of complete telomere dysfunction. Notably, *ctc1 tert* telomeres shorten more rapidly than in either single mutant background (K. Boltz and D. Shippen, unpublished data), arguing that telomerase continues to act on telomeres in the absence of CST.

Although the molecular basis for this ATR-independent pathway of DNA damage-induced telomerase repression is unknown, such a response reduces the potential for telomerase to act at sites of DNA damage, thereby limiting the chance of inappropriate telomere formation. A variety of mechanisms have been reported in yeast and vertebrates to restrain telomerase action following genotoxic stress (Schulz and Zakian, 1994; Kharbanda *et al.*, 2000; Wong *et al.*, 2002; Makovets and Blackburn, 2009). The extent to which all of these pathways are conserved bears further investigation. Finally, it is curious that despite the strong inhibition of telomerase in plants lacking ATR, telomere length homeostasis is unperturbed in the five generations of mutants we monitored (Vespa *et al.*, 2005); this study). One possibility is that DNA damage triggers a qualitative change in telomerase behavior, which is detected in our Q-TRAP assay as a quantitative change in activity. Repeat addition processivity (RAP) is not a property of *Arabidopsis* telomerase that can be assessed in our PCR-based TRAP assay. However, RAP of telomerase influences, and is influenced by, telomere

length (Lue, 2004). Telomerase RAP is dramatically altered in human cancer cells depending upon whether telomeres are within the normal range, or are artificially shortened (Zhao *et al.*, 2011). Likewise, the RAP of yeast telomerase is enhanced at critically shortened telomeres in an ATM-dependent manner (Chang *et al.*, 2007). Thus, it is conceivable that a crippled telomerase in *atr* mutants is sufficient to maintain telomeres already in the wild type range, but lacks the capacity to act efficiently on critically shortened telomeres in *ctc1* mutants, thereby enhancing the pace of telomere attrition.